HIGHER MATHEMATICS
FOR
ENGINEERS AND PHYSICISTS

The quality of the material used in the manufacture of this book is governed by continued postwar shortages.

Higher Mathematics

for

Engineers and Physicists

BY

IVAN S. SOKOLNIKOFF, Ph.D.
Professor of Mathematics, University of Wisconsin

AND

ELIZABETH S. SOKOLNIKOFF, Ph.D.
Lecturer in Mathematics, University of Wisconsin

SECOND EDITION
THIRTEENTH IMPRESSION

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK AND LONDON
1941

HIGHER MATHEMATICS FOR ENGINEERS AND PHYSICISTS

PRINTED IN THE UNITED STATES OF AMERICA

THE MAPLE PRESS COMPANY, YORK, PA.

PREFACE

The favorable reception of the First Edition of this volume appears to have sustained the authors' belief in the need of a book on mathematics beyond the calculus, written from the point of view of the student of applied science. The chief purpose of the book is to help to bridge the gap which separates many engineers from mathematics by giving them a bird's-eye view of those mathematical topics which are indispensable in the study of the physical sciences.

It has been a common complaint of engineers and physicists that the usual courses in advanced calculus and differential equations place insufficient emphasis on the art of formulating physical problems in mathematical terms. There may also be a measure of truth in the criticism that many students with pronounced utilitarian leanings are obliged to depend on books that are more distinguished for rigor than for robust uses of mathematics.

This book is an outgrowth of a course of lectures offered by one of the authors to students having a working knowledge of the elementary calculus. The keynote of the course is the practical utility of mathematics, and considerable effort has been made to select those topics which are of most frequent and immediate use in applied sciences and which can be given in a course of one hundred lectures. The illustrative material has been chosen for its value in emphasizing the underlying principles rather than for its direct application to specific problems that may confront a practicing engineer.

In preparing the revision the authors have been greatly aided by the reactions and suggestions of the users of this book in both academic and engineering circles. A considerable portion of the material contained in the First Edition has been rearranged and supplemented by further illustrative examples, proofs, and problems. The number of problems has been more than doubled. It was decided to omit the discussion of improper integrals and to absorb the chapter on Elliptic Integrals into

much enlarged chapters on Infinite Series and Differential Equations. A new chapter on Complex Variable incorporates some of the material that was formerly contained in the chapter on Conformal Representation. The original plan of making each chapter as nearly as possible an independent unit, in order to provide some flexibility and to enhance the availability of the book for reference purposes, has been retained.

I. S. S.
E. S. S.

MADISON, WISCONSIN,
September, 1941.

CONTENTS

HIGHER MATHEMATICS FOR ENGINEERS AND PHYSICISTS

CHAPTER I

INFINITE SERIES

It is difficult to conceive of a single mathematical topic that occupies a more prominent place in applied mathematics than the subject of infinite series. Students of applied sciences meet infinite series in most of the formulas they use, and it is quite essential that they acquire an intelligent understanding of the concepts underlying the subject.

The first section of this chapter is intended to bring into sharper focus some of the basic (and hence more difficult) notions with which the reader became acquainted in the first course in calculus. It is followed by ten sections that are devoted to a treatment of the algebra and calculus of series and that represent the minimum theoretical background necessary for an intelligent use of series. Some of the practical uses of infinite series are indicated briefly in the remainder of the chapter and more fully in Chaps. II, VII, and VIII.

1. Fundamental Concepts. Familiarity with the concepts discussed in this section is essential to an understanding of the contents of this chapter.

FUNCTION. *The variable y is said to be a function of the variable x if to every value of x under consideration there corresponds at least one value of y.*

If x is the variable to which values are assigned at will, then it is called the *independent variable.* If the values of the variable y are determined by the assignment of values to the independent

variable x, then y is called the *dependent variable*. The functional dependence of y upon x is usually denoted by the equation*

$$y = f(x).$$

Unless a statement to the contrary is made, it will be supposed in this book that the variable x is permitted to assume real values only and that the corresponding values of y are also real. In this event the function $f(x)$ is called a *real function of the real variable x*. It will be observed that

$$y = \sqrt{x} \tag{1-1}$$

does not represent a real function of x for all real values of x, for the values of y become imaginary if x is negative. In order that the symbol $f(x)$ define a real function of x, it may be necessary to restrict the range of values that x may assume. Thus, (1-1) defines a real function of x only if $x \geq 0$. On the other hand, $y = \sqrt{x^2 - 1}$ defines a real function of x only if $|x| \geq 1$.

Sequences and Limits. Let some process of construction yield a succession of values

$$x_1, x_2, x_3, \cdots, x_n, \cdots,$$

where it is assumed that every x_i is followed by other terms. Such a succession of terms is called an *infinite sequence*. Examples of sequences are

(a) $1, 2, 3, \cdots, n, \cdots,$

(b) $\frac{1}{2}, -\frac{1}{4}, \frac{1}{8}, -\frac{1}{16}, \cdots, (-1)^{n-1}\frac{1}{2^n}, \cdots,$

(c) $0, 2, 0, 2, \cdots, 1 + (-1)^n, \cdots.$

Sequences will be considered here only in connection with the theorems on infinite series,† and for this purpose it is necessary to have a definition of the limit of a sequence.

Definition. *The sequence $x_1, x_2, \cdots, x_n, \cdots$ is said to converge to the constant L as a limit if for any preassigned positive number ϵ, however small, one can find a positive integer p such that*

$$|x_n - L| < \epsilon \qquad \textit{for all } n > p.$$

* Other letters are often used. In particular, if more than one function enters into the discussion, the functions may be denoted by $f_1(x)$, $f_2(x)$, etc.; by $f(x)$, $g(x)$, etc.; by $F(x)$, $G(x)$, etc.

† For a somewhat more extensive treatment, see I. S. Sokolnikoff, Advanced Calculus, pp. 3–21.

For convenience, this definition* is frequently written in the compact form

$$\lim_{n\to\infty} x_n = L,$$

and L is called the *limit* of the sequence. If a variable x takes on these successive values $x_1, x_2, \cdots, x_n, \cdots$, then x is said to *approach L as a limit.* It follows from this definition that, of the sequences given above, (*b*) converges to the limit 0, whereas (*a*) and (*c*) are not convergent.

As an illustration, let the variable x assume the set of values

$$x_1 = 0.1, \qquad x_2 = 0.11, \qquad x_3 = 0.111, \cdots.$$

It is easily seen that

$$\lim_{n\to\infty} x_n = \frac{1}{9};$$

that is, corresponding to any $\epsilon > 0$, one can find a positive integer p such that

$$|L - x_n| \equiv \left|\frac{1}{9} - x_n\right| < \epsilon$$

for all values of n greater than p. Observe that

$$\frac{1}{9} - x_1 = \frac{1}{90}, \qquad \frac{1}{9} - x_2 = \frac{1}{900}, \quad \cdots, \qquad \frac{1}{9} - x_n = \frac{1}{9\cdot 10^n}.$$

Hence, for any ϵ that is chosen, it is necessary to demand that n be large enough so that

$$\frac{1}{9} - x_n \equiv \frac{1}{9\cdot 10^n} < \epsilon.$$

The inequality is equivalent to

$$9\cdot 10^n > \frac{1}{\epsilon},$$

and, taking logarithms to the base 10,*

$$\log 9 + n > \log\frac{1}{\epsilon}$$

or

$$n > -(\log 9 + \log \epsilon) = -\log 9\epsilon.$$

* From the definition of the logarithm, it follows that, if $A > B$, then $\log A > \log B$.

Thus, if p is chosen as any integer greater than $|-\log 9\epsilon|$, the inequality

$$\left|\frac{1}{9} - x_n\right| < \epsilon$$

will be satisfied for all values of n greater than p.

INFINITE SERIES. Let $u_1, u_2, u_3, \cdots$ be an infinite sequence of real functions of a real variable x. Then the symbol

$$\sum_{n=1}^{\infty} u_n(x) \equiv u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots \tag{1-2}$$

is called an *infinite series*.

If, in (1-2), x is assigned some fixed value, say $x = x_0$, there results the series of constants

$$\sum_{n=1}^{\infty} u_n(x_0). \tag{1-3}$$

Denote by $s_n(x_0)$ the nth partial sum, that is, the sum of the first n terms, of the series (1-3) so that

$$s_n(x_0) = u_1(x_0) + u_2(x_0) + \cdots + u_n(x_0).$$

As n increases indefinitely, the sequence of constants

$$s_1(x_0),\ s_2(x_0),\ \cdots,\ s_n(x_0),\ \cdots$$

either will converge to a finite limit S or it will not converge to such a limit. If

$$\lim_{n\to\infty} s_n(x_0) = S,$$

the series (1-2) is said to converge to the value S for $x = x_0$.* If the series (1-2) converges for every value of x in some interval† (a, b), then the series is said to be convergent in the interval (a, b).

As an example, consider the series

$$1 + x + x^2 + \cdots + x^{n-1} + \cdots. \tag{1-4}$$

If $x = \frac{1}{2}$, (1-4) becomes

$$1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{n-1}} + \cdots,$$

* This limit S is usually called the *sum of the series* (1-3).

† This means that x can assume any real value between a and b and that a and b can be thought of as the end points of an interval of the x-axis.

which is convergent to the value 2. In order to establish this fact, note that

$$s_n = 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{n-1}}$$

is a geometric progression of ratio $\frac{1}{2}$, so that

$$s_n = \frac{1 - \dfrac{1}{2^n}}{1 - \dfrac{1}{2}} = 2 - \frac{1}{2^{n-1}}.$$

Hence, the absolute value of the difference between 2 and s_n is $1/2^{n-1}$, which can be made arbitrarily small by choosing n sufficiently large.

On the other hand, if $x = -1$, the series (1-4) becomes

$$1 - 1 + 1 - 1 + \cdots + (-1)^{n-1} + \cdots,$$

which does not converge; for $s_{2n} = 0$ and $s_{2n-1} = 1$ for any choice of n and, therefore, $\lim\limits_{n\to\infty} s_n$ does not exist. Moreover, if $x = 2$, the series (1-4) becomes

$$1 + 2 + 4 + \cdots + 2^{n-1} + \cdots,$$

so that s_n increases indefinitely with n and $\lim\limits_{n\to\infty} s_n$ does not exist. If an infinite series does not converge for a certain value of x, it is said to *diverge* or *be divergent* for that value of x. It will be shown later that the series (1-4) is convergent for $-1 < x < 1$ and divergent for all other values of x.

The definition of the limit, as given above, assumes that the value of the limit S is known. Frequently it is possible to infer the existence of S without actually knowing its value. The following example will serve to illustrate this point.

Example. Consider the series

$$s = 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots,$$

and compare the sum of its first n terms

$$s_n = 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!}$$

with the sum of the geometrical progression

$$S_n = 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}}$$
$$= 2 - \frac{1}{2^{n-1}}.$$

The corresponding terms of S_n are never less than those of s_n; but, no matter how large n be taken, S_n is less than 2. Consequently, $s_n < 2$; and since the successive values of s_n form an increasing sequence of numbers, the sum of the first series must be greater than 1 and less than or equal to 2. A geometrical interpretation of this statement may help to fix the idea. If the successive values of s_n,

$$s_1 = 1,$$
$$s_2 = 1 + \frac{1}{2!} = 1.5,$$
$$s_3 = 1 + \frac{1}{2!} + \frac{1}{3!} = 1.667,$$
$$s_4 = 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} = 1.708,$$
$$s_5 = 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} = 1.717,$$

are plotted as points on a straight line (Fig. 1), the points representing the sequence $s_1, s_2, \cdots, s_n, \cdots$ always move to the right but never

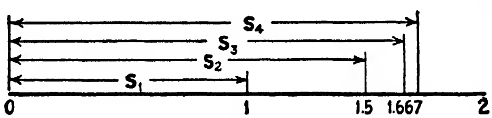

FIG. 1.

progress as far as the point 2. It is intuitively clear that there must be some point s, either lying to the left of 2 or else coinciding with it, which the numbers s_n approach as a limit. In this case the numerical value of the limit has not been ascertained, but its existence was established with the aid of what is known as the *fundamental principle.*

Stated in precise form the principle reads as follows: *If an infinite set of numbers $s_1, s_2, \cdots, s_n, \cdots$ forms an increasing sequence* (that is, *$s_N > s_n$, when $N > n$) and is such that every s_n is less than some fixed number M* (that is, *$s_n < M$ for all values of n), then s_n approaches a limit s that is not greater than M* (that is, $\lim_{n \to \infty} s_n = s \leq M$). The formulation of the principle for a decreasing sequence of numbers $s_1, s_2, \cdots, s_n, \cdots$, which are always greater than a certain fixed number m, will be left to the reader.

2. Series of Constants. The definition of the convergence of a series of functions evidently depends on a study of the behavior

of series of constants. The reader has had some acquaintance with such series in his earlier study of mathematics, but it seems desirable to provide a summary of some essential theorems that will be needed later in this chapter. The following important theorem gives the necessary and sufficient condition for the convergence of an infinite series of constants:

THEOREM. *The infinite series of constants* $\sum_{n=1}^{\infty} u_n$ *converges if and only if there exists a positive integer* n *such that for all positive integral values of* p

$$|s_{n+p} - s_n| \equiv |u_{n+1} + u_{n+2} + \cdots + u_{n+p}| < \epsilon,$$

where ϵ *is any preassigned positive constant.*

The necessity of the condition can be proved immediately by recalling the definition of convergence. Thus, assume that the series converges, and let its sum be S, so that

$$\lim_{n\to\infty} s_n = S$$

and also, for any fixed value of p,

$$\lim_{n\to\infty} s_{n+p} = S.$$

Hence,

$$\lim_{n\to\infty} (s_{n+p} - s_n) \equiv \lim_{n\to\infty} (u_{n+1} + u_{n+2} + \cdots + u_{n+p}) = 0,$$

which is another way of saying that

$$|u_{n+1} + u_{n+2} + \cdots + u_{n+p}| < \epsilon$$

for a sufficiently large value of n.

The proof of the sufficiency of the condition requires a fair degree of mathematical maturity and will not be given here.*

This theorem is of great theoretical importance in a variety of investigations, but it is seldom used in any practical problem requiring the testing of a given series. A number of tests for convergence, applicable to special types of series, will be given in the following sections.

It may be remarked that a sufficient condition that a series diverge is that the terms u_n do not approach zero as a limit when n increases indefinitely. Thus the necessary condition for convergence of a series is that $\lim_{n\to\infty} u_n = 0$, but this condition is not

* See SOKOLNIKOFF, I. S., Advanced Calculus, pp. 11–13.

sufficient; that is, there are series for which $\lim_{n\to\infty} u_n = 0$ but which are not convergent. A classical example illustrating this case is the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots,$$

in which S_n increases without limit as n increases.

Despite the fact that a proof of the divergence of the harmonic series is given in every good course in elementary calculus, it will be recalled here because of its importance in subsequent considerations. Since

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{n+n} > n \cdot \frac{1}{2n} = \frac{1}{2},$$

it is possible, beginning with any term of the series, to add a definite number of terms and obtain a sum greater than ½. If $n = 2$,

$$\frac{1}{3} + \frac{1}{4} > \frac{1}{2};$$

$n = 4$,

$$\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > \frac{1}{2};$$

$n = 8$,

$$\frac{1}{9} + \frac{1}{10} + \cdots + \frac{1}{16} > \frac{1}{2};$$

$n = 16$,

$$\frac{1}{17} + \frac{1}{18} + \cdots + \frac{1}{32} > \frac{1}{2}.$$

Thus it is possible to group the terms of the harmonic series

$$1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \cdots$$

in such a way that the sum of the terms in each parenthesis exceeds ½; and, since the series

$$1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots$$

is obviously divergent, the harmonic series is divergent also.

3. Series of Positive Terms. This section is concerned with series of the type

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \cdots + a_n + \cdots,$$

where the a_n are positive constants. It is evident from the definition of convergence and from the fundamental principle (see Sec. 1) that the convergence of a series of positive constants will be established if it is possible to demonstrate that the partial sums s_n remain bounded. This means that there exists some positive number M such that $s_n < M$ for all values of n. The proof of the following important test is based on such a demonstration.

COMPARISON TEST. *Let $\sum_{n=1}^{\infty} a_n$ be a series of positive terms, and let $\sum_{n=1}^{\infty} b_n$ be a series of positive terms that is known to converge. Then the series $\sum_{n=1}^{\infty} a_n$ is convergent if there exists an integer p such that, for $n \geq p$, $a_n \leq b_n$. On the other hand, if $\sum_{n=1}^{\infty} c_n$ is a series of positive terms that is known to be divergent and if $a_n \geq c_n$ for $n \geq p$, then $\sum_{n=1}^{\infty} a_n$ is divergent also.*

Since the convergence or divergence of a series evidently is not affected by the addition or subtraction of a finite number of terms, the proof will be given on the assumption that $p = 1$. Let $s_n = a_1 + a_2 + \cdots + a_n$, and let B denote the sum of the series $\sum_{n=1}^{\infty} b_n$ and B_n its nth partial sum. Then, since $a_n \leq b_n$ for all values of n, it follows that $s_n \leq B_n$ for all values of n. Hence, the s_n remain bounded, and the series $\sum_{n=1}^{\infty} a_n$ is convergent. On the other hand, if $a_n \geq c_n$ for all values of n and if the series $\sum_{n=1}^{\infty} c_n$ diverges, then the series $\sum_{n=1}^{\infty} a_n$ will diverge also.

There are two series that are frequently used as series for comparison.

a. The geometric series

$$a + ar + ar^2 + \cdots + ar^n + \cdots, \tag{3-1}$$

which the reader will recall* is convergent to $\dfrac{a}{1-r}$ if $|r| < 1$ and is divergent if $|r| \geq 1$.

b. The p series

$$1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} + \cdots, \tag{3-2}$$

which converges if $p > 1$ and diverges if $p \leq 1$.

Consider first the case when $p > 1$, and write (3-2) in the form

$$1 + \left(\frac{1}{2^p} + \frac{1}{3^p}\right) + \left(\frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p}\right) + \cdots + \left(\frac{1}{(2^{n-1})^p} + \cdots + \frac{1}{(2^n - 1)^p}\right) + \cdots, \tag{3-3}$$

where the nth term of (3-3) contains 2^{n-1} terms of the series (3-2). Each term, after the first, of (3-3) is less than the corresponding term of the series

$$1 + 2 \cdot \frac{1}{2^p} + 4 \cdot \frac{1}{4^p} + \cdots + 2^{n-1} \frac{1}{(2^{n-1})^p} + \cdots,$$

or

$$1 + \frac{1}{2^{p-1}} + \frac{1}{(2^{p-1})^2} + \cdots + \frac{1}{(2^{p-1})^{n-1}} + \cdots. \tag{3-4}$$

Since the geometric series (3-4) has a ratio $1/2^{p-1}$ (which is less than unity for $p > 1$), it is convergent and, by the comparison test, (3-2) will converge also.

If $p = 1$, (3-2) becomes the harmonic series which has been shown to be divergent.

If $p < 1$, $1/n^p > 1/n$ for $n > 1$, so that each term of (3-2), after the first, is greater than the corresponding term of the harmonic series; hence, the series (3-2) is divergent also.

Example 1. Test the series

$$1 + \frac{1}{2^2} + \frac{1}{3^3} + \cdots + \frac{1}{n^n} + \cdots.$$

The geometric series

$$1 + \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^n} + \cdots$$

* Since the sum of the geometric progression of n terms $a + ar + ar^2 + \cdots + ar^{n-1}$ is equal to $\dfrac{a - ar^n}{1 - r} = \dfrac{a}{1 - r}(1 - r^n)$.

is known to be convergent, and the terms of the geometric series are never less than the corresponding terms of the given series. Hence, the given series is convergent.

Example 2. Test the series

$$1 + \frac{1}{\log 2} + \frac{1}{\log 3} + \frac{1}{\log 4} + \cdots + \frac{1}{\log n} + \cdots.$$

Compare the terms of this series with the terms of the p series for $p = 1$,

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n} + \cdots.$$

The given series is divergent, for its terms (after the first) are greater than the corresponding terms of the p series, which diverges when $p = 1$.

Ratio Test. *The series $\sum_{n=1}^{\infty} a_n$ of positive terms is convergent if*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = r < 1$$

and divergent if

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} > 1.$$

If $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 1$, the series may converge or diverge.

Consider first the case when $r < 1$, and let q denote some constant between r and 1. Then there will be some positive integer N such that

$$\frac{a_{n+1}}{a_n} < q \qquad \text{for all } n \geq N.$$

Hence,

$$\begin{aligned} a_{N+1} &< a_N q, \\ a_{N+2} &< a_{N+1} q < a_N q^2, \\ a_{N+3} &< a_{N+2} q < a_N q^3, \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

and

$$a_{N+1} + a_{N+2} + a_{N+3} + \cdots < a_N(q + q^2 + q^3 + \cdots).$$

Since $q < 1$, the series in the right-hand member is convergent; therefore, the series in the left-hand member converges, also. It follows that the series $\sum_{n=1}^{\infty} a_n$ is convergent.

If the limit of the ratio is greater than 1, then $a_{n+1} > a_n$ for every $n \geq N$ so that $\lim_{n\to\infty} a_n \neq 0$, and hence the series $\sum_{n=1}^{\infty} a_n$ is divergent.

It is important to observe that this theorem makes no reference to the magnitude of the ratio of a_{n+1}/a_n but deals solely with the *limit* of the ratio. Thus, in the case of the harmonic series the ratio is $a_{n+1}/a_n = n/(n + 1)$, which remains less than 1 for all finite values of n, but the limit of the ratio is precisely equal to 1. Hence the test gives no information in this case.

Example 1. For the series

$$1 + \frac{2}{2} + \frac{3}{2^2} + \frac{4}{2^3} + \cdots + \frac{n}{2^{n-1}} + \cdots,$$

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{n+1}{2^n} \cdot \frac{2^{n-1}}{n} = \lim_{n\to\infty} \frac{1 + \frac{1}{n}}{2} = \frac{1}{2}$$

and, therefore, the series converges.

Example 2. The series

$$\frac{1}{10} + \frac{2!}{10^2} + \frac{3!}{10^3} + \cdots + \frac{n!}{10^n} + \cdots$$

is divergent, for

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{(n+1)!}{10^{n+1}} \frac{10^n}{n!} = \lim_{n\to\infty} \frac{n+1}{10} = \infty.$$

Example 3. Test the series

$$\frac{1}{1\cdot 2} + \frac{1}{3\cdot 4} + \frac{1}{5\cdot 6} + \cdots + \frac{1}{(2n-1)2n} + \cdots.$$

Here

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{1}{(2n+1)(2n+2)} \cdot \frac{(2n-1)2n}{1}$$

$$= \lim_{n\to\infty} \frac{4n^2 - 2n}{4n^2 + 6n + 2} = \lim_{n\to\infty} \frac{1 - \frac{1}{2n}}{1 + \frac{3}{2n} + \frac{1}{2n^2}} = 1.$$

Hence, the test fails; but if the given series be compared with the p series for $p = 2$, it is seen to be convergent.

CAUCHY'S INTEGRAL TEST. *Let* $\sum_{n=1}^{\infty} a_n$ *be a series of positive terms such that* $a_{n+1} < a_n$. *If there exists a positive decreasing function* $f(x)$, *for* $x \geq 1$, *such that* $f(n) = a_n$, *then the given series converges if the integral*

$$\int_1^\infty f(x)\, dx$$

exists; the series diverges if the integral does not exist.

The proof of this test is deduced easily from the following graphical considerations. Each term a_n of the series may be thought of as representing the area of a rectangle of base unity and height $f(n)$ (see Fig. 2). The sum of the areas of the first n inscribed rectangles is less than $\int_1^{n+1} f(x)\,dx$, so that

$$s_{n+1} - a_1 \equiv a_2 + a_3 + \cdots + a_{n+1} < \int_1^{n+1} f(x)\,dx.$$

But $f(x)$ is positive, and hence

$$\int_1^{n+1} f(x)\,dx < \int_1^{\infty} f(x)\,dx.$$

If the integral on the right exists, it follows that the partial sums are bounded and, therefore, the series converges (see Sec. 1).

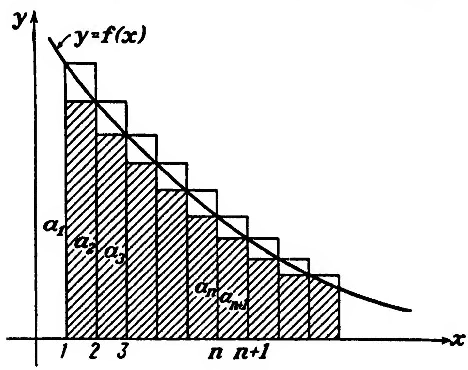

FIG. 2.

The sum of the areas of the circumscribed rectangles, $a_1 + a_2 + \cdots + a_n$, is greater than $\int_1^{n+1} f(x)\,dx$; hence, the series will diverge if the integral does not exist.

Example 1. Test the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots.$$

In this case, $f(x) = \dfrac{1}{x}$ and

$$\int_1^{\infty} \frac{1}{x}\,dx \equiv \lim_{n\to\infty} \int_1^n \frac{dx}{x} = \lim_{n\to\infty} \log n = \infty,$$

and the series is divergent.

Example 2. Apply Cauchy's test to the p series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ where $p > 0$. Taking $f(x) = \frac{1}{x^p}$, observe that

$$\int_1^n \frac{dx}{x^p} = \frac{1}{1-p} x^{1-p}\Big|_1^n, \qquad \text{if } p \neq 1,$$
$$= \log x\Big|_1^n, \qquad \text{if } p = 1.$$

Hence, $\int_1^{\infty} \frac{dx}{x^p}$ exists if $p > 1$ and does not exist if $p \leq 1$.

PROBLEMS

1. Test for convergence

(*a*) $\frac{1}{2} + \frac{1}{2 \cdot 2^2} + \frac{1}{3 \cdot 2^3} + \frac{1}{4 \cdot 2^4} + \cdots;$

(*b*) $\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \cdots + \frac{1}{2n} + \cdots;$

(*c*) $1 + \frac{2!}{2^2} + \frac{3!}{3^2} + \cdots;$

(*d*) $\frac{1}{3} + \frac{1 \cdot 2}{3 \cdot 5} + \frac{1 \cdot 2 \cdot 3}{3 \cdot 5 \cdot 7} + \cdots;$

(*e*) $\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots;$

(*f*) $1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \cdots;$

(*g*) $\frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \cdots;$

(*h*) $\frac{1}{2 \log 2} + \frac{1}{3 \log 3} + \frac{1}{4 \log 4} + \cdots.$

2. Use Cauchy's integral test to investigate the convergence of

(*a*) $\frac{1}{2} + \frac{1}{5} + \frac{1}{10} + \cdots + \frac{1}{1 + n^2} + \cdots;$

(*b*) $1 + \frac{2}{1 + 2^2} + \frac{3}{2 + 3^2} + \cdots.$

3. Show that the series $\sum_{n=1}^{\infty} a_n$ of positive terms is divergent if na_n has a limit L which is different from 0. *Hint:* Let $\lim_{n \to \infty} na_n = L$ so that $na_n > L - \epsilon$ for n large enough. Hence, $a_n > \frac{L - \epsilon}{n}$.

4. Test for convergence

(a) $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$;

(b) $\sum_{n=1}^{\infty} \frac{n}{2^n}$;

(c) $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n(n+1)}}$;

(d) $\sum_{n=1}^{\infty} \frac{n!}{n^2}$;

(e) $\sum_{n=1}^{\infty} \frac{1}{n \cdot 2^n}$;

(f) $\sum_{n=1}^{\infty} \frac{n^2}{n!}$;

(g) $\sum_{n=1}^{\infty} \frac{n!}{10^n}$;

(h) $\sum_{n=1}^{\infty} \frac{1}{n\sqrt{n}}$;

(i) $\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2}$;

(j) $\sum_{n=0}^{\infty} \frac{n}{(2n+1)^2}$.

4. Alternating Series. A series whose terms are alternately positive and negative is called an *alternating series*. There is a simple test, due to Leibnitz, that establishes the convergence of many of these series.

TEST FOR AN ALTERNATING SERIES. *If the alternating series $a_1 - a_2 + a_3 - a_4 + \cdots$, where the a_i are positive, is such that $a_{n+1} < a_n$ and $\lim_{n\to\infty} a_n = 0$, then the series is convergent. Moreover, if S is the sum of the series, the numerical value of the difference between S and the nth partial sum is less than a_{n+1}.*

Since

$$\begin{aligned} s_{2n} &= (a_1 - a_2) + (a_3 - a_4) + \cdots + (a_{2n-1} - a_{2n}) \\ &= a_1 - (a_2 - a_3) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n}, \end{aligned}$$

it is evident that s_{2n} is positive and also that $s_{2n} < a_1$ for all values of n. Also, $s_2 < s_4 < s_6 < \cdots$, so that these partial

sums tend to a limit S (by the fundamental principle). Since $s_{2n+1} = s_{2n} + a_{2n+1}$ and $\lim_{n\to\infty} a_{2n+1} = 0$, it follows that the partial sums of odd order tend to this same limit. Therefore, the series converges. The proof of the second statement of the test will be left as an exercise for the reader.

Example 1. The series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

is convergent since $\lim_{n\to\infty} \frac{1}{n} = 0$ and $\frac{1}{n+1} < \frac{1}{n}\cdot$ Moreover, $s_4 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4}$ differs from the sum S by less than $\frac{1}{5}$.

Example 2. The series

$$1 - \frac{1}{2} + \frac{1}{3^1} - \frac{1}{4} + \frac{1}{3^2} - \frac{1}{6} + \frac{1}{3^3} - \cdots$$

is divergent. Why?

5. Series of Positive and Negative Terms. The alternating series and the series of positive constants are special types of the general series of constants in which the terms can be either positive or negative.

DEFINITION. *If* $u_1 + u_2 + \cdots + u_n + \cdots$ *is an infinite series of terms such that the series of the absolute values of its terms,* $|u_1| + |u_2| + \cdots + |u_n| + \cdots$, *is convergent, then the series* $u_1 + u_2 + \cdots + u_n + \cdots$ *is said to be absolutely convergent. If the series of absolute values is not convergent, but the given series is convergent, then the given series is said to be conditionally convergent.*

Thus,

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

is convergent, but the series of absolute values,

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots,$$

is not, so that the original series is conditionally convergent.

If a series is absolutely convergent, it can be shown that the series formed by changing the signs of any of the terms is also a convergent series. This is an immediate result of the following theorem:

Theorem. *If the series of absolute values* $\sum_{n=1}^{\infty} |u_n|$ *is convergent, then the series* $\sum_{n=1}^{\infty} u_n$ *is necessarily convergent.*

Let

$$s_n = u_1 + u_2 + \cdots + u_n$$

and

$$t_n = |u_1| + |u_2| + \cdots + |u_n|.$$

If p_n denotes the sum of the positive terms occurring in s_n and $-q_n$ denotes the sum of the negative terms, then

$$s_n = p_n - q_n \tag{5-1}$$

and

$$t_n = p_n + q_n.$$

The series $\sum_{n=1}^{\infty} |u_n|$ is assumed to be convergent, so that

$$\lim_{n\to\infty} t_n \equiv \lim_{n\to\infty} (p_n + q_n) \equiv L. \tag{5-2}$$

But p_n and q_n are positive and increasing with n and, since (5-2) shows that both remain less than L, it follows from the fundamental principle that both the p_n and q_n sequences converge. If

$$\lim_{n\to\infty} p_n = P \qquad \text{and} \qquad \lim_{n\to\infty} q_n = Q,$$

then (5-1) gives

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} (p_n - q_n) = P - Q,$$

which establishes the convergence of $\sum_{n=1}^{\infty} u_n$.

Moreover, it can be shown that changing the order of the terms in an absolutely convergent series gives a series which is convergent to the same value as the original series.* However, conditionally convergent series do not possess this property. In fact, by suitably rearranging the order of the terms of a conditionally convergent series, the resulting series can be made to converge to any desired value. For example, it is known† that the sum of the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} + \cdots$$

* See Sokolnikoff, I. S., Advanced Calculus, pp. 240–241.

† See Example 1, Sec. 13.

is $\log_e 2$. The fact that the sum of this series is less than 1 and greater than ½ can be made evident by writing the series as

$$\left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{5} - \frac{1}{6}\right) + \cdots,$$

which shows that the value of $s_n > \frac{1}{2}$ for $n > 2$; whereas, by writing it as

$$1 - \left(\frac{1}{2} - \frac{1}{3}\right) - \left(\frac{1}{4} - \frac{1}{5}\right) \cdots,$$

it is clear that $s_n < 1$ for $n \geq 2$. Some questions might be raised concerning the legitimacy of introducing parentheses in a convergent infinite series. The fact that the associative law holds unrestrictedly for convergent infinite series can be established easily directly from the definition of the sum of the infinite series. It will be shown* next that it is possible to rearrange the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

so as to obtain a new series whose sum is equal to 1. The positive terms of this series in their original order are

$$1, \quad \frac{1}{3}, \quad \frac{1}{5}, \quad \frac{1}{7}, \quad \frac{1}{9}, \cdots.$$

The negative terms are

$$-\frac{1}{2}, \quad -\frac{1}{4}, \quad -\frac{1}{6}, \quad -\frac{1}{8}, \cdots.$$

In order to form a series that converges to 1, first pick out, in order, as many positive terms as are needed to make their sum equal to or just greater than 1, then pick out just enough negative terms so that the sum of all terms so far chosen will be just less than 1, then more positive terms until the sum is just greater than 1, etc. Thus, the partial sums will be

$$s_1 = 1,$$

$$s_2 = 1 - \frac{1}{2} = \frac{1}{2},$$

$$s_4 = 1 - \frac{1}{2} + \frac{1}{3} + \frac{1}{5} = \frac{31}{30},$$

* General proof can be constructed along the lines of this example.

$$s_5 = 1 - \frac{1}{2} + \frac{1}{3} + \frac{1}{5} - \frac{1}{4} = \frac{47}{60},$$

$$s_7 = 1 - \frac{1}{2} + \frac{1}{3} + \frac{1}{5} - \frac{1}{4} + \frac{1}{7} + \frac{1}{9} = \frac{1307}{1260},$$

$$s_8 = 1 - \frac{1}{2} + \frac{1}{3} + \frac{1}{5} - \frac{1}{4} + \frac{1}{7} + \frac{1}{9} - \frac{1}{6} = \frac{1093}{1260},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

It is clear that the series formed by this method will have a sum equal to 1.

As another example, consider the conditionally convergent series

$$1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \cdots. \tag{5-3}$$

Let the order of the terms in (5-3) be rearranged to give the series

$$\left(1 + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{2}}\right) + \left(\frac{1}{\sqrt{5}} + \frac{1}{\sqrt{7}} - \frac{1}{\sqrt{4}}\right) + \left(\frac{1}{\sqrt{9}} + \frac{1}{\sqrt{11}} - \frac{1}{\sqrt{6}}\right) + \cdots. \tag{5-4}$$

The nth term of (5-4) is

$$a_n \equiv \frac{1}{\sqrt{4n-3}} + \frac{1}{\sqrt{4n-1}} - \frac{1}{\sqrt{2n}},$$

which is greater than

$$b_n \equiv \frac{1}{\sqrt{4n}} + \frac{1}{\sqrt{4n}} - \frac{1}{\sqrt{2n}} = \left(1 - \frac{1}{\sqrt{2}}\right)\frac{1}{\sqrt{n}}.$$

But the series $\sum_{n=1}^{\infty} b_n$ is divergent, and it follows that the series (5-4) must diverge.

Inasmuch as the series $\sum_{n=1}^{\infty} |u_n|$ is a series of positive terms, the tests that were developed in Sec. 3 can be applied in establishing the absolute convergence of the series $\sum_{n=1}^{\infty} u_n$. In particular, the ratio test can be restated in the following form:

RATIO TEST. *The series* $\sum_{n=1}^{\infty} u_n$ *is absolutely convergent if*

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| < 1$$

and is divergent if

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| > 1.$$

If the limit is unity, the test gives no information.

Example 1. In the case of the series

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots,$$

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{x^n}{n!}\frac{(n-1)!}{x^{n-1}}\right| = \lim_{n\to\infty}\left|\frac{x}{n}\right| = 0$$

for all values of x. Hence, the series is convergent for all values of x and, in particular, the series

$$1 - 2 + \frac{2^2}{2!} - \frac{2^3}{3!} + \cdots$$

is absolutely convergent.

Example 2. Consider the series

$$\frac{1}{1-x} + \frac{1}{2(1-x)^2} + \frac{1}{3(1-x)^3} + \cdots.$$

Here

$$\lim_{n\to\infty}\left|\frac{1}{(n+1)(1-x)^{n+1}}\cdot\frac{n(1-x)^n}{1}\right| = \lim_{n\to\infty}\left|\frac{n}{(n+1)(1-x)}\right|$$
$$= \lim_{n\to\infty}\left|\frac{1}{\left(1+\frac{1}{n}\right)(1-x)}\right| = \frac{1}{|1-x|}.$$

Therefore, the series will converge if

$$\frac{1}{|1-x|} < 1 \qquad \text{or} \qquad 1 < |1-x|,$$

which is true for $x < 0$ and for $x > 2$.

For $x = 0$ and for $x = 2$ the limit is unity, but if $x = 0$ the series becomes the divergent harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots,$$

and if $x = 2$ there results the convergent alternating series

$$-1 + \frac{1}{2} - \frac{1}{3} + \cdots + (-1)^n \frac{1}{n} + \cdots .$$

It follows that the original series converges for $x < 0$ and for $x \geq 2$ and diverges for $0 \leq x < 2$.

6. Algebra of Series. The following important theorems are stated without proof:*

THEOREM 1. *Any two convergent series*

$$U = u_1 + u_2 + \cdots + u_n + \cdots$$
$$V = v_1 + v_2 + \cdots + v_n + \cdots$$

can be added or subtracted term by term to give

$$U + V = (u_1 + v_1) + (u_2 + v_2) + \cdots + (u_n + v_n) + \cdots$$

or

$$U - V = (u_1 - v_1) + (u_2 - v_2) + \cdots + (u_n - v_n) + \cdots .$$

If the original series are both absolutely convergent, then the resulting series will be absolutely convergent also.

THEOREM 2. *If*

$$U = u_1 + u_2 + \cdots + u_n + \cdots$$

and

$$V = v_1 + v_2 + \cdots + v_n + \cdots$$

are two absolutely convergent series, then they can be multiplied like finite sums and the product series will converge to UV. Moreover, the product series will be absolutely convergent. Thus,

$$UV = u_1v_1 + u_1v_2 + u_2v_1 + u_1v_3 + u_2v_2 + u_3v_1 + \cdots .$$

THEOREM 3. *In an absolutely convergent series the positive terms by themselves form a convergent series and also the negative terms by themselves form a convergent series. If in a convergent series the positive terms form a divergent series, then the series of negative terms is also divergent and the original series is conditionally convergent.*

THEOREM 4. *If $u_1 + u_2 + \cdots + u_n + \cdots$ is an absolutely convergent series and if $M_1, M_2, \cdots, M_n, \cdots$ is any sequence of quantities whose numerical values are all less than some positive number N, then the series*

$$u_1M_1 + u_2M_2 + \cdots + u_nM_n + \cdots$$

is absolutely convergent.

* See SOKOLNIKOFF, I. S., Advanced Calculus, pp. 212–213, 241–245.

Example. Consider the series

$$\frac{\sin x}{1^3} - \frac{\sin 2x}{2^3} + \frac{\sin 3x}{3^3} - \cdots.$$

This series is absolutely convergent for all values of x, for the series

$$\frac{1}{1^3} - \frac{1}{2^3} + \frac{1}{3^3} - \cdots$$

is absolutely convergent and $|\sin nx| \leq 1$.

PROBLEMS

1. Show that the following series are divergent:

(a) $\frac{5}{2} - \frac{7}{4} + \frac{9}{6} - \frac{11}{8} + \cdots + (-1)^{n-1}\frac{2n+3}{2n} + \cdots;$

(b) $\frac{1}{2} - \frac{1}{2^2} + \frac{1}{3} - \frac{1}{3^2} + \frac{1}{4} - \frac{1}{4^2} + \cdots;$

(c) $\frac{2}{1} - \frac{3}{2} + \frac{4}{3} - \frac{5}{4} + \cdots.$

2. Test for convergence, and if the series is convergent determine whether it is absolutely convergent.

(a) $1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \cdots;$

(b) $\frac{1}{3} - \frac{1\cdot 3}{3\cdot 6} + \frac{1\cdot 3\cdot 5}{3\cdot 6\cdot 9} - \frac{1\cdot 3\cdot 5\cdot 7}{3\cdot 6\cdot 9\cdot 12} + \cdots;$

(c) $\frac{2}{1} - \frac{3}{2} + \frac{4}{3} - \frac{5}{4} + \cdots.$

3. For what values of x are the following series convergent?

(a) $x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1}\frac{x^n}{n} + \cdots;$

(b) $1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots;$

(c) $1 - x + x^2 - x^3 + \cdots;$

(d) $\frac{1}{x} + \frac{1}{2x^2} + \frac{1}{3x^3} + \cdots + \frac{1}{nx^n} + \cdots.$

4. Determine the intervals of convergence of the following series:

(a) $\frac{2x}{x+4} + \frac{1}{2}\left(\frac{2x}{x+4}\right)^2 + \frac{1}{3}\left(\frac{2x}{x+4}\right)^3 + \cdots;$

(b) $x + 2!x^2 + 3!x^3 + 4!x^4 + \cdots;$

(c) $1 + mx + \frac{m(m-1)}{2!}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \cdots.$

where m is not a positive integer.

7. Continuity of Functions. Uniform Convergence. Before proceeding with a discussion of infinite series of functions, it is necessary to have a clear understanding of the concept of continuity of functions. The reader will recall that a function $f(x)$ is said to be continuous at a point $x = x_0$ if $\lim_{x \to x_0} f(x) = f(x_0)$ regardless of how x approaches x_0. From the discussion of the limit in Sec. 1, it appears that this concept can be defined in the following way:

DEFINITION. *The function $f(x)$ is continuous at the point $x = x_0$ if, corresponding to any preassigned positive number ϵ, it is possible to find a positive number δ such that*

$$|f(x) - f(x_0)| < \epsilon \textit{ whenever } |x - x_0| < \delta. \tag{7-1}$$

The foregoing analytical definition of continuity is merely a formulation in exact mathematical language of the intuitive

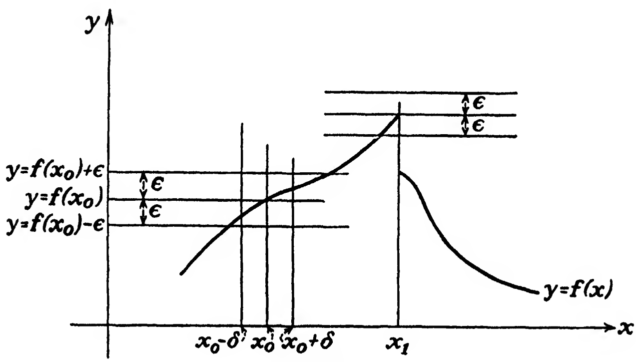

FIG. 3.

concept of continuity. If the function $f(x)$ is represented by a graph and if it is continuous at the point $x = x_0$, then it is possible to find a strip bounded by the two parallel lines $x = x_0 + \delta$ and $x = x_0 - \delta$, such that the graph of the function will lie entirely within the strip bounded by the parallel lines $y = f(x_0) + \epsilon$ and $y = f(x_0) - \epsilon$ (Fig. 3). But if the function is discontinuous at some point (such as $x = x_1$), then no interval about such a point can be found such that the graph of the function will lie entirely within the strip of width 2ϵ, where ϵ is arbitrarily small.

DEFINITION. *A function is said to be continuous in an interval (a, b) if it is continuous at each point of the interval.*

If a finite number of functions that are all continuous in an interval (a, b) are added together, the sum also will be a continuous function in (a, b). The question arises as to whether this property will be retained in the case of an infinite series of continuous functions. Moreover, it is frequently desirable to obtain the derivative (or integral) of a function $f(x)$ by means of term-by-term differentiation (or integration) of an infinite series that defines $f(x)$. Unfortunately, such operations are not always valid, and many important investigations have led to erroneous results solely because of the improper handling of infinite series. A discussion of such questions requires an introduction of the property of uniform convergence of a series.

It was stated in Sec. 1 that the series

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots \tag{7-2}$$

is convergent to the value S, when $x = x_0$, provided that

$$\lim_{n\to\infty} s_n(x_0) = S. \tag{7-3}$$

The statement embodied in (7-3) means that for any preassigned positive number ϵ, however small, one can find a positive number N such that

$$|s_n(x_0) - S| < \epsilon \qquad \text{for all } n \geq N.$$

If the series (7-2) is convergent for every value of x in the interval (a, b), then the series (7-2) defines a function $S(x)$. Let x_0 be some value of x in (a, b), so that

$$|s_n(x_0) - S(x_0)| < \epsilon \quad \text{whenever } n \geq N.$$

It is important to note that, in general, the magnitude of N depends not only on the choice of ϵ, but also on the value of x_0.

This last remark may be clarified by considering the series

$$x + (x - 1)x + (x - 1)x^2 + \cdots + (x - 1)x^{n-1} + \cdots. \tag{7-4}$$

Since

$$s_n(x) = x + (x - 1)x + (x - 1)x^2 + \cdots + (x - 1)x^{n-1} = x^n,$$

it is evident that

$$\lim_{n\to\infty} s_n(x) = \lim_{n\to\infty} x^n = 0, \quad \text{if } 0 \leq x < 1.$$

Thus, $S(x) = 0$ for all values of x in the interval $0 \leq x < 1$, and therefore

$$|s_n(x) - S(x)| = |x^n - 0| = |x^n|.$$

Hence, the requirement that $|s_n(x) - S(x)| < \epsilon$, for an arbitrary ϵ, will be satisfied only if $x^n < \epsilon$. This inequality leads to the condition

$$n \log x < \log \epsilon.$$

Since $\log x$ is negative for x between 0 and 1, it follows that it is necessary to have

$$n > \frac{\log \epsilon}{\log x},$$

which clearly shows the dependence of N on both ϵ and x. In fact, if $\epsilon = 0.01$ and $x = 0.1$, n must be greater than $\log 0.01/\log 0.1 = -2/-1 = 2$, so that N can be chosen as any number greater than 2. If $\epsilon = 0.01$ and $x = 0.5$, N must be chosen larger than $\log 0.01/\log 0.5$, which is greater than 6. Since the values of $\log x$ approach zero as x approaches unity, it appears that the ratio $\log \epsilon/\log x$ will increase indefinitely and that it will be impossible to find a single value of N which will serve for $\epsilon = 0.01$ and for all values of x in $0 \leq x < 1$.

It should be noted that the discussion applies to the interval $(0, 1)$ and that it might be possible to find an N, depending on ϵ only, if some other interval were chosen. If the series and the interval are such that it is possible to find an N, for any preassigned ϵ, which will serve *for all values of x in the interval,* then the series is said to converge *uniformly* in the interval.

DEFINITION OF UNIFORM CONVERGENCE. *The series* $\sum_{n=1}^{\infty} u_n(x)$ *is uniformly convergent in the interval* (a, b) *if, for any* $\epsilon > 0$, *there exists a positive number* N, *independent of the value of* x *in* (a, b), *such that*

$$|S(x) - s_n(x)| < \epsilon \qquad \textit{for all } n \geq N.$$

The distinction between uniform convergence and the type of convergence exemplified by the discussion of the series (7-4) will become apparent in the discussion of the series

$$1 + x + x^2 + \cdots + x^n + \cdots, \tag{7-5}$$

where $-\frac{1}{2} \leq x \leq \frac{1}{2}$.

Since $s_n(x) = \dfrac{1 - x^n}{1 - x}$, it follows that

$$S(x) = \lim_{n\to\infty} s_n(x) = \lim_{n\to\infty} \left(\frac{1}{1 - x} - \frac{x^n}{1 - x}\right) = \frac{1}{1 - x}.$$

Then,

$$|S(x) - s_n(x)| = \left|\frac{x^n}{1 - x}\right|,$$

which will be less than an arbitrary $\epsilon > 0$ if

$$|x^n| < \epsilon(1 - x).$$

Hence,

$$n \log |x| < \log \epsilon(1 - x),$$

or

$$n > \frac{\log \epsilon(1 - x)}{\log |x|}. \tag{7-6}$$

Again, it appears that the choice of N will depend on both x and ϵ, but in this case it is possible to choose an N that will serve for all values of x in $(-\frac{1}{2}, \frac{1}{2})$. Observing that the ratio $\log \epsilon(1 - x)/\log |x|$ assumes its maximum value, for a fixed ϵ, when $x = -\frac{1}{2}$, it is evident that if N is chosen so that

$$N > \frac{\log \epsilon(\frac{3}{2})}{\log \frac{1}{2}} = 1 - \frac{\log 3\epsilon}{\log 2},$$

then the inequality (7-6) will be satisfied for all $n \geqq N$.

Upon recalling the conditions for uniform convergence, it is seen that the series (7-5) converges uniformly for $-\frac{1}{2} \leq x \leq \frac{1}{2}$. However, it should be noted that (7-5) does not converge uniformly in the interval $(-1, 1)$. For, in this interval, the ratio appearing in (7-6) will increase indefinitely as x approaches the values ± 1. The discussion given above shows that the series (7-5) is uniformly convergent in any interval $(-a, a)$, where $a < 1$.

It may be remarked that the series (7-5) does not even converge for $x = \pm 1$. For $x = 1$, it is obviously divergent, and when $x = -1$ the series becomes

$$1 - 1 + 1 - 1 + \cdots.$$

If $-1 < x < 1$, (7-5) defines the function $\dfrac{1}{1 - x}$, which takes the value $\frac{1}{2}$ when $x = -1$.

As is often the case with definitions, the definition of uniform convergence is usually difficult to apply when the behavior of a particular series is to be investigated. There are available several tests for the uniform convergence of series, the simplest of which is associated with the name of the German mathematician Weierstrass.

THEOREM. (WEIERSTRASS M TEST). *Let*

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots \tag{7-7}$$

be a series of functions of x defined in the interval (a, b). If there exists a convergent series of positive constants,

$$M_1 + M_2 + \cdots + M_n + \cdots,$$

such that $|u_i(x)| \leq M_i$ for all values of x in (a, b), then the series (7-7) is uniformly and absolutely convergent in (a, b).

Since, by hypothesis, the series of M's is convergent, it follows that for any prescribed $\epsilon > 0$ there exists an N such that

$$M_{n+1} + M_{n+2} + \cdots < \epsilon \quad \text{for all } n \geq N.$$

But $|u_i(x)| \leq M_i$ for all values of x in (a, b), so that

$$|u_{n+1}(x) + u_{n+2}(x) + \cdots| \leq M_{n+1} + M_{n+2} + \cdots < \epsilon$$

for all $n \geq N$ and for all values of x in (a, b). Therefore, the series (7-7) is uniformly and absolutely convergent in (a, b).

The fact that the Weierstrass test establishes the absolute convergence, as well as the uniform convergence, of a series means that it is applicable only to series which converge absolutely. There are other tests that are not so restricted, but these tests are more complex. It should be emphasized that a series may converge uniformly but not absolutely, and vice versa.

Example 1. Consider the series

$$\frac{\sin x}{1^2} + \frac{\sin 2x}{2^2} + \cdots + \frac{\sin nx}{n^2} + \cdots.$$

Since $|\sin nx| \leq 1$ for all values of x, the convergent series

$$\frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2} + \cdots$$

will serve as an M series. It follows that the given series is uniformly and absolutely convergent in any interval, no matter how large.

Example 2. As noted earlier in this section, the series

$$1 + x + x^2 + \cdots + x^n + \cdots$$

converges uniformly in any interval $(-a, a)$, where $a < 1$. The series of positive constants

$$1 + a + a^2 + \cdots + a^n + \cdots$$

could be used as an M series in this case, since this series converges for $a < 1$ and $|x^i| \leq a^i$ for x in $(-a, a)$.

PROBLEMS

1. Show that the series (7-4) is uniformly convergent in the interval $(0, \frac{1}{2})$.

2. By using the definition of uniform convergence, show that

$$\frac{1}{x+1} - \frac{1}{(x+1)(x+2)} - \cdots - \frac{1}{(x+n-1)(x+n)} - \cdots$$

is uniformly convergent in the interval $0 \leq x \leq 1$.

Hint: Rewrite the series to show that $s_n(x) = \dfrac{1}{x+n}$ and therefore

$$|S(x) - s_n(x)| = \frac{1}{x+n}.$$

3. Test the following series for uniform convergence:

(a) $\dfrac{\cos x}{1^2} + \dfrac{\cos 3x}{3^2} + \dfrac{\cos 5x}{5^2} + \cdots;$

(b) $\dfrac{\sin 2x}{1 \cdot 3} + \dfrac{\sin 4x}{3 \cdot 5} + \dfrac{\sin 6x}{5 \cdot 7} + \cdots;$

(c) $1 + x \cos \theta + x^2 \cos 2\theta + x^3 \cos 3\theta + \cdots, \ |x| \leq x_1 < 1;$

(d) $\dfrac{\cos 2x}{2^\alpha} - \dfrac{\cos 3x}{3^\alpha} + \dfrac{\cos 4x}{4^\alpha} - \cdots;$

(e) $10x + 10^2x^2 + 10^3x^3 + \cdots.$

8. Properties of Uniformly Convergent Series. As remarked in the preceding section, the concept of uniform convergence was introduced in order to allow the discussion of certain properties of infinite series. This section contains the statements* of three important theorems concerning uniformly convergent series.

THEOREM 1. *Let*

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots$$

be a series such that each $u_i(x)$ is a continuous function of x in the interval (a, b). If the series is uniformly convergent in (a, b), then the sum of the series is also a continuous function of x in (a, b).

* For proofs, see I. S. Sokolnikoff, Advanced Calculus, pp. 256–262.

COROLLARY. *A discontinuous function cannot be represented by a uniformly convergent series of continuous functions in the neighborhood of the point of discontinuity.*

THEOREM 2. *If a series of continuous functions,*

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots,$$

converges uniformly to $S(x)$ *in* (a, b), *then*

$$\int_\alpha^\beta S(x)\,dx = \int_\alpha^\beta u_1(x)\,dx + \int_\alpha^\beta u_2(x)\,dx + \cdots + \int_\alpha^\beta u_n(x)\,dx + \cdots,$$

where $a < \alpha < b$ *and* $a < \beta < b$.

THEOREM 3. *Let*

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots$$

be a series of differentiable functions that converges to $S(x)$ *in* (a, b). *If the series*

$$u_1'(x) + u_2'(x) + \cdots + u_n'(x) + \cdots$$

converges uniformly in (a, b), *then it converges to* $S'(x)$.

These theorems provide sufficient conditions only. It may be that the sum of the series is a continuous function when the series is not uniformly convergent. It is impossible to discuss necessary conditions in this brief introduction to uniform convergence. It may happen also that the series is differentiable or integrable term by term when it does not converge uniformly. In the chapter on Fourier series it will be shown that a discontinuous function can be represented by an infinite series of continuous functions. In that chapter, it is established that the series

$$2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots\right)$$

represents the function x for $-\pi < x < \pi$. But, if this series be differentiated term by term, the resulting series is

$$2(\cos x - \cos 2x + \cos 3x - \cdots),$$

which does not converge in $(-\pi, \pi)$; for the necessary condition for convergence, namely, that $\lim_{n\to\infty} |u_n| = 0$, does not hold for any value of x.

The series used in the first example of Sec. 7,

$$\text{(8-1)} \qquad \frac{\sin x}{1^2} + \frac{\sin 2x}{2^2} + \frac{\sin 3x}{3^2} + \cdots + \frac{\sin nx}{n^2} + \cdots,$$

is uniformly convergent in any interval (a, b) and as such defines a continuous function $S(x)$. Moreover, the series can be integrated term by term to produce the integral of $S(x)$. The term-by-term derivative of (8-1) is

$$\cos x + \tfrac{1}{2}\cos 2x + \tfrac{1}{3}\cos 3x + \cdots, \tag{8-2}$$

which is convergent in $(0, \pi)$, but the M series for (8-2) cannot be found since $1 + \frac{1}{2} + \frac{1}{3} + \cdots$ is divergent. This merely suggests that (8-2) may not converge to the derivative of $S(x)$, but it does not say that it will not.

PROBLEMS

1. Test for uniform convergence the series obtained by term-by-term differentiation of the five series given in Prob. 3 of Sec. 7.

2. Test for uniform convergence the series obtained by term-by-term integration of the five series given in Prob. 3, Sec. 7.

9. Power Series. One of the most important types of infinite series of functions is the power series

$$\sum_{n=0}^{\infty} a_n x^n \equiv a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots, \tag{9-1}$$

in which the a_i are independent of x. Some of the reasons for the usefulness of power series will become apparent in the discussion that follows.

Whenever a series of functions is used, the first question which arises is that of determining the values of the variable for which the series is convergent. The ratio test was applied for this purpose in the examples discussed in Sec. 5. In general, for a power series,

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{a_n}{a_{n-1}}\,x\right|,$$

so that the series converges if

$$\lim_{n\to\infty}\left|\frac{a_n}{a_{n-1}}\,x\right| < 1$$

and diverges if

$$\lim_{n\to\infty}\left|\frac{a_n}{a_{n-1}}\,x\right| > 1.$$

Therefore, the series will converge for those values of x for which

$$|x| < \lim_{n\to\infty}\left|\frac{a_{n-1}}{a_n}\right|.$$

If $\lim\limits_{n\to\infty}\left|\dfrac{a_{n-1}}{a_n}\right| = r$, it follows that the series will converge when x lies inside the interval $(-r, r)$, which is called the *interval of convergence,* the number r being called the *radius of convergence.* This discussion establishes the following theorem:

THEOREM. *If the series* $\sum\limits_{n=0}^{\infty} a_n x^n$ *is such that*

$$\lim_{n\to\infty}\left|\frac{a_{n-1}}{a_n}\right| = r,$$

then the series converges in the interval $-r < x < r$ *and diverges outside this interval. The series may or may not converge at the end points of the interval.*

Example 1. Consider the series

$$1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^n}{n} + \cdots.$$

Since $\lim\limits_{n\to\infty}\left|\dfrac{a_{n-1}}{a_n}\right| = \lim\limits_{n\to\infty}\left|\dfrac{n}{n-1}\right| = 1$, the series converges for $-1 < x < 1$ and diverges for $|x| > 1$. At the end point $x = -1$ the series becomes

$$1 - 1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \cdots$$

which is convergent. At the end point $x = 1$ the divergent series

$$1 + 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$$

is obtained. Hence, this power series is convergent for $-1 \leq x < 1$.

Example 2. The series

$$1 + x + 2!x^2 + \cdots + n!x^n + \cdots$$

will serve to demonstrate the fact that there are power series which converge only for the value $x = 0$. For

$$\lim_{n\to\infty}\left|\frac{a_{n-1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{(n-1)!}{n!}\right| = \lim_{n\to\infty}\frac{1}{n} = 0.$$

Obviously, the series converges for $x = 0$, as does every power series, but it diverges for every other value of x.

Power series in $x - h$ are frequently more useful than the special case in which the value of h is zero. A series of this type has the form

$$a_0 + a_1(x-h) + a_2(x-h)^2 + \cdots + a_n(x-h)^n + \cdots.$$

In this case the test ratio yields

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{a_n}{a_{n-1}}\right| |x-h|.$$

If this limit is less than 1, the series is convergent; if greater than 1, the series is divergent; and if the limit is equal to 1, the test fails and the values of x, which make the limit equal to 1, must be investigated. Thus, if the series is

$$1 + (x-1) + \frac{(x-1)^2}{2^2} + \frac{(x-1)^3}{3^2} + \cdots + \frac{(x-1)^n}{n^2} + \cdots,$$

then

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{(x-1)^n}{n^2}\,\frac{(n-1)^2}{(x-1)^{n-1}}\right|$$
$$= \lim_{n\to\infty}\left(1-\frac{2}{n}-\frac{1}{n^2}\right)|x-1| = |x-1|.$$

Therefore the series converges if $|x-1| < 1$, or $0 < x < 2$, and diverges for $|x-1| > 1$, or $x < 0$, $x > 2$. For $x - 1 = 1$, or $x = 2$, the series becomes

$$1 + 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} + \cdots,$$

which is the p series for $p = 2$ and is therefore convergent. For $x - 1 = -1$, or $x = 0$, the series becomes

$$1 - 1 + \frac{1}{2^2} - \frac{1}{3^2} + \cdots + (-1)^n\frac{1}{n^2} + \cdots,$$

which is an alternating series of decreasing terms with $\lim_{n\to\infty} u_n = 0$ and is therefore convergent. Thus the series is convergent for $0 \le x \le 2$.

PROBLEM

Find the interval of convergence for each of the following series, and determine its behavior at the end points of the interval:

(*a*) $1 + x + x^2 + x^3 + \cdots;$

(*b*) $1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots;$

(*c*) $1 + (x+1) + \frac{(x+1)^2}{2} + \frac{(x+1)^3}{3} + \cdots;$

(d) $1 - 2x + 3x^2 - 4x^3 + \cdots$;

(e) $1 + \frac{x}{2} + \frac{x^2}{2^2 \cdot 2} + \frac{x^3}{2^3 \cdot 3} + \cdots$;

(f) $(x - 2) - \frac{1}{2}(x - 2)^2 + \frac{1}{3}(x - 2)^3 - \cdots$;

(g) $x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$.

10. Properties of Power Series. The importance of power series in applied mathematics is due to the properties given in the theorems of this section, as will be evident from the applications discussed in succeeding sections.

THEOREM 1. *If $r > 0$ is the radius of convergence of a power series $\sum_{n=0}^{\infty} a_n x^n$, then the series converges absolutely and uniformly for every value of x in any interval $a \leq x \leq b$ that is interior to $(-r, r)$.*

Since the interval (a, b) lies entirely within the interval $(-r, r)$, it is possible to choose a positive number c that is less than r but greater than a and b. The interval (a, b) will then lie entirely within the interval $(-c, c)$; and it follows that, for $a \leq x \leq b$,

$$|a_n x^n| < |a_n c^n|.$$

The series of positive constants $\sum_{n=0}^{\infty} |a_n c^n|$ is convergent, for $c < r$, and, accordingly, can be used as a Weierstrass M series establishing the absolute and uniform convergence of $\sum_{n=0}^{\infty} a_n x^n$ in (a, b).

THEOREM 2. *A power series $\sum_{n=0}^{\infty} a_n x^n$ defines a continuous function for all values of x in any closed interval (a, b) that is interior to the interval of convergence of the series.*

This statement is a direct consequence of the preceding theorem and of Theorem 1, Sec. 8.

THEOREM 3. *If the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n x^n$ is r, then the radii of convergence of the series $\sum_{n=0}^{\infty} n a_n x^{n-1}$ and $\sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}$, obtained by term-by-term differentiation and integration of the given series, are also r.*

If the radius of convergence can be determined from the ratio test, then the proof follows immediately from the fact that if

$$\lim_{n\to\infty}\left|\frac{a_{n-1}}{a_n}\right| = r, \qquad \text{then} \qquad \lim_{n\to\infty}\left|\frac{(n-1)a_{n-1}}{na_n}\right| = r \qquad \text{and}$$

$$\lim_{n\to\infty}\left|\frac{(n+1)a_{n-1}}{na_n}\right| = r.$$

Since the series obtained by term-by-term differentiation and integration are also power series, these processes can be repeated as many times as desired and the resulting series will be power series that converge in the interval $(-r, r)$. It follows from Theorem 1 that all these series are uniformly and absolutely convergent in any interval which is interior to $(-r, r)$. However, the behavior of these series at the end points $x = -r$ and $x = r$ must be investigated in each case.

For example, the series

$$1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^n}{n} + \cdots$$

has unity for its radius of convergence. The series converges for $x = -1$ but is divergent for $x = 1$. The series obtained by term-by-term differentiation is

$$1 + x + x^2 + \cdots + x^n + \cdots,$$

which has the same radius of convergence but diverges at both $x = 1$ and $x = -1$. On the other hand, the series obtained by term-by-term integration is

$$x + \frac{x^2}{1\cdot 2} + \frac{x^3}{2\cdot 3} + \frac{x^4}{3\cdot 4} + \cdots + \frac{x^{n+1}}{n(n+1)} + \cdots$$

which converges for both $x = 1$ and $x = -1$.

This discussion leads to the conclusion stated in the following theorem:

THEOREM 4. *A power series* $\sum_{n=0}^{\infty} a_n x^n$ *may be differentiated and integrated term by term as many times as desired in any closed interval* (a, b) *that lies entirely within the interval of convergence of the given series.*

THEOREM 5. *If a power series* $\sum_{n=0}^{\infty} a_n x^n$ *vanishes for all values of* x *lying in a certain interval about the point* $x = 0$, *then the*

coefficient of each power of x vanishes, that is,

$$a_0 = 0, \qquad a_1 = 0, \qquad a_2 = 0, \cdots, \qquad a_n = 0, \cdots.$$

The reader may attempt to construct the proof of this theorem with the aid of Theorem 2 of this section.

11. Expansion of Functions in Power Series. It was stated in Theorem 2, Sec. 10, that a power series defines a continuous function of x in any interval which lies within the interval of convergence. This theorem suggests at once the possibility of using such a power series for the purpose of computation. For example, the values of $\sin x$ might be obtained by means of a power series. Accordingly, it becomes necessary to develop some method of obtaining such a power series, and this section is devoted to a derivation of Taylor's formula and a discussion of Taylor's series.

One of the simplest proofs of Taylor's formula will be given here.* It assumes that the given function $f(x)$ has a continuous nth derivative throughout the interval (a, b). Taylor's formula is obtained by integrating this nth derivative n times in succession between the limits a and x, where x is any point in (a, b). Thus,

$$\int_a^x f^{(n)}(x)\,dx = f^{(n-1)}(x)\Big|_a^x = f^{(n-1)}(x) - f^{(n-1)}(a),$$

$$\int_a^x \int_a^x f^{(n)}(x)\,(dx)^2 = \int_a^x f^{(n-1)}(x)\,dx - \int_a^x f^{(n-1)}(a)\,dx$$
$$= f^{(n-2)}(x) - f^{(n-2)}(a) - (x-a)f^{(n-1)}(a),$$

$$\int_a^x \int_a^x \int_a^x f^{(n)}(x)\,(dx)^3 = f^{(n-3)}(x) - f^{(n-3)}(a) - (x-a)f^{(n-2)}(a)$$
$$- \frac{(x-a)^2}{2!}f^{(n-1)}(a),$$

. . . .

$$\int_a^x \cdots \int_a^x f^{(n)}(x)\,(dx)^n = f(x) - f(a) - (x-a)f'(a)$$
$$- \frac{(x-a)^2}{2!}f''(a) - \cdots - \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a).$$

* For other proofs, see I. S. Sokolnikoff, Advanced Calculus, pp. 291–295.

Solving for $f(x)$ gives

$$f(x) = f(a) + (x - a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + R_n, \tag{11-1}$$

where

$$R_n = \int_a^x \cdots \int_a^x f^{(n)}(x)\,(dx)^n. \tag{11-2}$$

The formula given by (11-1) is known as Taylor's formula and the particular form of R_n given in (11-2) is called the integral form of the remainder after n terms. The foregoing can be stated in the form of a theorem.

TAYLOR'S THEOREM. *Any function $f(x)$ that possesses a continuous derivative $f^{(n)}(x)$ in the interval (a, b) can be expanded in the form* (11-1) *for all values of x in (a, b).*

The term R_n, which represents the difference between $f(x)$ and the polynomial of degree $n - 1$ in $x - a$, is frequently more useful when expressed in a different form. Since*

$$\int_a^x f^{(n)}(x)\,dx = (x - a)f^{(n)}(\xi), \qquad \text{where } a < \xi < x,$$

repeated integration gives

$$R_n = \int_a^x \cdots \int_a^x f^{(n)}(x)\,(dx)^n = \frac{(x-a)^n}{n!}f^{(n)}(\xi). \tag{11-3}$$

The right-hand member of (11-3) is the Lagrangian form of the remainder after n terms.

The special form of Taylor's formula that is obtained by setting $a = 0$ is known as the Maclaurin formula. In this case

$$f(x) = f(0) + f'(0)x + f''(0)\frac{x^2}{2!} + \cdots + f^{(n-1)}(0)\frac{x^{n-1}}{(n-1)!} + R_n, \tag{11-4}$$

where

$$R_n = f^{(n)}(\xi)\frac{x^n}{n!}, \qquad 0 < \xi < x.$$

* The student will recall from elementary calculus that

$$\int_a^b \varphi(x)\,dx = (b - a)\varphi(\xi), \qquad \text{where } a < \xi < b.$$

Taylor's formula with the Lagrangian form of the remainder is often encountered in a somewhat different form, which results from setting $x - a = h$. Since $a < \xi < x$, ξ can be written in the form $a + \theta h$, where $0 < \theta < 1$. Hence, (11-1) becomes

$$(11\text{-}5)\quad f(a + h) = f(a) + f'(a)h + f''(a)\frac{h^2}{2!} + \cdots + f^{(n-1)}(a)\frac{h^{n-1}}{(n-1)!} + f^{(n)}(a + \theta h)\frac{h^n}{n!}, \qquad \text{where } 0 < \theta < 1.$$

In this derivation of Taylor's formula, it was assumed that $f(x)$ possesses a continuous nth derivative, and as a result it appeared that then $f(x)$ could be expressed as a polynomial of degree n in $x - a$. It should be noted, however, that only the first n coefficients of this polynomial are constants, for the coefficient of $(x - a)^n$ is a function of ξ and the value of ξ is dependent upon the choice of x. It may happen that $f(x)$ possesses derivatives of all orders and that the remainder R_n approaches zero as a limit when $n \to \infty$ regardless of the choice of x in (a, b). If such is the case, the infinite series

$$(11\text{-}6)\quad f(a) + f'(a)(x - a) + f''(a)\frac{(x-a)^2}{2!} + \cdots + f^{(n)}(a)\frac{(x-a)^n}{n!} + \cdots$$

is convergent and, in general,* it converges to $f(x)$.

The series given in (11-6) is called the Taylor's series expansion, or representation, of the function $f(x)$ about the point $x = a$. The special form of (11-6) that is obtained when $a = 0$, namely,

$$(11\text{-}7)\quad f(0) + f'(0)x + f''(0)\frac{x^2}{2!} + \cdots + f^{(n)}(0)\frac{x^n}{n!} + \cdots$$

is called Maclaurin's series.

Example. Find the Taylor's series expansion of $\cos x$ in powers of $x - \frac{\pi}{2}$.

Since

$$f(x) = \cos x, \qquad f\left(\frac{\pi}{2}\right) = 0;$$

$$f'(x) = -\sin x, \qquad f'\left(\frac{\pi}{2}\right) = -1;$$

* For a further discussion of this point, see I. S. Sokolnikoff, Advanced Calculus, pp. 296–298.

$$f''(x) = -\cos x, \qquad f''\left(\frac{\pi}{2}\right) = 0;$$

$$f'''(x) = \sin x, \qquad f'''\left(\frac{\pi}{2}\right) = 1;$$

$$f^{\text{IV}}(x) = \cos x, \qquad f^{\text{IV}}\left(\frac{\pi}{2}\right) = 0;$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots;$$

it follows that the result is

$$\cos x = -\left(x - \frac{\pi}{2}\right) + \frac{1}{3!}\left(x - \frac{\pi}{2}\right)^3 - \frac{1}{5!}\left(x - \frac{\pi}{2}\right)^5 + \cdots.$$

Since it is often possible to obtain a power series expansion of a function $f(x)$ by some other method, the question arises as to the relation of such an expansion to the Taylor's series expansion for $f(x)$. For example, a power series expansion for the function $\frac{1}{1-x}$ is obtained easily by division, giving

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + \cdots.$$

The reader can check the fact that the Maclaurin expansion for this function is identical with the power series obtained by division. That this is not an exceptional case is established in the following theorem:

Uniqueness Theorem. *There is only one possible expansion of a function in a power series in $x - a$; and, therefore, if such an expansion be found in any manner whatsoever, it must coincide with Taylor's expansion about the point a.*

Suppose that $f(x)$ could be represented by two power series in $x - a$, so that

$$f(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \cdots + a_n(x - a)^n + \cdots$$

and

$$f(x) = b_0 + b_1(x - a) + b_2(x - a)^2 + \cdots + b_n(x - a)^n + \cdots.$$

Since both these expansions represent $f(x)$ in the vicinity of a, there must be some interval about the point $x = a$ in which both the expansions are valid. Then, in this interval,

$$\sum_{n=0}^{\infty} a_n(x-a)^n = \sum_{n=0}^{\infty} b_n(x-a)^n,$$

or

$$\sum_{n=0}^{\infty} (a_n - b_n)(x-a)^n = 0.$$

It follows from Theorem 5, Sec. 10, that

$$a_n - b_n = 0, \qquad (n = 0, 1, 2, \cdots),$$

or

$$a_n = b_n, \qquad (n = 0, 1, 2, \cdots).$$

Hence, the two power series expansions are identical.

Taylor's formula is frequently more useful in a slightly modified form. Let

$$x - a \equiv h,$$

so that

$$x = a + h.$$

Then

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + \frac{f^{(n)}(\xi)}{n!}(x-a)^n$$

becomes

$$f(a+h) = f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}h^{n-1} + \frac{f^{(n)}(a+\theta h)}{n!}h^n, \tag{11-8}$$

in which $0 < \theta < 1$, so that $a < a + \theta h < a + h$.

PROBLEMS

1. Find the expansion of each of the following functions in power series in x:

(a) e^x, (b) $\sin x$, (c) $\cos x$, (d) $\tan^{-1} x$,
(e) $\sin^{-1} x$, (f) $\sec x$, (g) $\tan x$, (h) $e^{\sin x}$.

2. Expand

(a) $\log x$ in powers of $x - 1$;

(b) $\tan x$ in powers of $x - \frac{\pi}{4}$;

(c) e^x in powers of $x - 2$;

(d) $\sin x$ in powers of $x - \frac{\pi}{6}$;

(e) $2 + x^2 - 3x^5 + 7x^6$ in powers of $x - 1$.

3. Show that $\sin x$ can be developed about any point a in a series (11-8) which converges for all values of h.

4. Differentiate term by term the power series in x for $\sin x$ and thus obtain the power series in x for $\cos x$. What is the interval of convergence of the resulting series?

5. Divide the series $\sin x \equiv x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$ by the series $\cos x \equiv 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$, and thus obtain the series for $\tan x$.

6. Differentiate the series for $\sin^{-1} x$ to obtain the expansion in powers of x for $(1 - x^2)^{-\frac{1}{2}}$. Find the interval of convergence. Is convergence absolute? Investigate the behavior of the series at the end points of the interval of convergence.

7. Establish with the aid of Maclaurin's series that

$$(a + b)^m \equiv k(1 + x)^m = k\left[1 + mx + \frac{m(m-1)}{2!}x^2 + \cdots\right],$$

where m is not a positive integer.

This series is convergent for $|x| < 1$ and divergent when $|x| > 1$. A complete discussion of this series will be found in Sokolnikoff's Advanced Calculus. Some facts are:

If $x = 1$, convergence is absolute if $m > 0$;
If $x = 1$, convergence is conditional if $0 > m > -1$;
If $x = -1$, convergence is absolute if $m > 0$;
If $x = -1$, series diverges when $m < 0$;
If $x = 1$, series diverges when $m \leq -1$.

8. Let $f(y) = \sum_{n=0}^{\infty} b_n y^n$ and $y = \sum_{n=0}^{\infty} a_n x^n$ be convergent power series. If $f(y)$ is a polynomial, then the powers of y in terms of x can be determined by repeated multiplications and thus the expansion for $f(y)$ in powers of x can be obtained. But if $f(y)$ is an infinite series, this procedure may not be valid. Inasmuch as the power series in x is always convergent for $x = 0$ and since the value of y for $x = 0$ is a_0, it is clear that the interval of convergence of $\sum_{n=0}^{\infty} b_n y^n$ must include a_0 if the series for $f(y)$, in powers of x, is to converge. But if $a_0 = 0$, then $f(y)$ surely can be expanded in power series in x by this method, for the point 0 is contained in the interval of convergence of $\sum_{n=0}^{\infty} b_n y^n$.

Apply this method to deriving the series in powers of x for $e^{\sin x}$ by setting

$$y = \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots.$$

and

$$e^y = 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \cdots.$$

Explain why this method fails to produce the series in powers of x for $\log(1 + e^x)$, where $e^x = y$.

12. Application of Taylor's Formula. In this section two illustrations of the application of Taylor's formula will be given, and in each case the remainder will be investigated to determine the error made in using the sum of the first n terms of the expansion instead of the function itself.

1. *Calculate the Value of* $\sin 10°$. Since $10°$ is closer to $0°$ than to any other value of x for which the values of $\sin x$ and its derivatives are known, the Maclaurin expansion for $\sin x$ will be determined and evaluated for $x = 10° = \pi/18$ radian. Then

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!}x^{n-1} + \frac{f^{(n)}(\xi)}{n!}x^n,$$

where $0 < \xi < \dfrac{\pi}{18}$.

Since

$$\begin{aligned}
f(x) &= \sin x, & f(0) &= 0;\\
f'(x) &= \cos x, & f'(0) &= 1;\\
f''(x) &= -\sin x, & f''(0) &= 0;\\
f'''(x) &= -\cos x, & f'''(0) &= -1;\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots;\\
f^{(n)}(x) &= \sin\left(x + \frac{n\pi}{2}\right), & f^{(n)}(0) &= \sin\frac{n\pi}{2};
\end{aligned}$$

therefore,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + \frac{x^n}{n!}f^{(n)}(\xi).$$

Here,

$$\begin{aligned}
R_n(x) \equiv \frac{x^n}{n!}f^{(n)}(\xi) &= \frac{x^n}{n!}f^{(n)}(\theta x), \qquad 0 < \theta < 1,\\
&= \frac{x^n}{n!}\sin\left(\theta x + \frac{n\pi}{2}\right).
\end{aligned}$$

If only the terms through x^7 (or x^8) are used in computing $\sin \pi/18$, the error will be

$$R_9\left(\frac{\pi}{18}\right) = \left(\frac{\pi}{18}\right)^9 \frac{1}{9!} \sin\left(\theta\frac{\pi}{18} + \frac{9}{2}\pi\right) = \left(\frac{\pi}{18}\right)^9 \frac{1}{9!} \cos\frac{\theta\pi}{18}$$
$$< \left(\frac{\pi}{18}\right)^9 \frac{1}{9!},$$

so that

$$\sin\frac{\pi}{18} = \frac{\pi}{18} - \left(\frac{\pi}{18}\right)^3 \frac{1}{3!} + \left(\frac{\pi}{18}\right)^5 \frac{1}{5!} - \left(\frac{\pi}{18}\right)^7 \frac{1}{7!},$$

with an error less than $\left(\frac{\pi}{18}\right)^9 \frac{1}{9!}$.

2. *Compute the Value of* $e^{1.1}$. It can be established readily, by expanding e^x in Maclaurin's series and evaluating for $x = 1$, that $e = 2.71828 \cdots$. In order to compute the value of $e^{1.1}$, the expansion of e^x about $x = 1$ will be used. The expansion is

$$f(x) = f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \cdots + \frac{f^{(n)}(\xi)}{n!}(x-1)^n.$$

Since

$$\begin{aligned} f(x) &= e^x, & f(1) &= e; \\ f'(x) &= e^x, & f'(1) &= e; \\ &\cdots\cdots & &\cdots\cdots \\ f^{(n)}(x) &= e^x, & f^{(n)}(1) &= e; \end{aligned}$$

and

$$f^{(n)}(\xi) = e^\xi, \qquad 1 < \xi < x;$$

therefore

$$e^x = e + e(x-1) + \frac{e}{2!}(x-1)^2 + \cdots + \frac{e}{(n-1)!}(x-1)^{n-1} + \frac{e^\xi}{n!}(x-1)^n.$$

Here

$$R_n = \frac{e^\xi}{n!}(x-1)^n,$$

so that the error made in using only four terms is

$$R_4 = \frac{e^\xi}{4!}(x-1)^4.$$

If $x = 1.1$,

$$e^{1.1} = e\left[1 + 0.1 + \frac{(0.1)^2}{2!} + \frac{(0.1)^3}{3!}\right] + \frac{e^\xi}{4!}(0.1)^4$$
$$= 1.105166e + \frac{0.0001}{24}e^\xi.$$

Thus, $e^{1.1} = 1.105166e$ with an error of $(0.0001/24)e^{\xi}$. Since ξ lies between 1 and 1.1 and since e^{ξ} is an increasing function, e^{ξ} is certainly less than e^2. An approximate value of e^2 is 7, and the error is certainly less than $0.0007/24 = 0.00003$. Therefore,

$$e^{1.1} = 1.1052e,$$

correct to four decimal places.

13. Evaluation of Definite Integrals by Means of Power Series. One of the most important applications of infinite series is their use in computing numerical values of definite integrals, such as $\int_0^1 e^{-x^2}\,dx$, in which the indefinite integral cannot be found in closed form. Moreover, the values of many transcendental functions are computed most easily by this method. Several examples of this use of infinite series are given in this section.

Example 1. Consider

$$\log(1+x) = \int_0^x \frac{dz}{1+z} = \int_0^x (1+z)^{-1}\,dz.$$

Since

$$(1+z)^{-1} = 1 - z + z^2 - z^3 + \cdots$$

for $|z| < 1$, it follows that

$$\begin{aligned}\log(1+x) &= \int_0^x \frac{dz}{1+z} = \int_0^x dz - \int_0^x z\,dz + \int_0^x z^2\,dz - \cdots \\ &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots .\end{aligned}$$

Example 2. Since

$$\begin{aligned}\sin^{-1} x &= \int_0^x \frac{dz}{\sqrt{1-z^2}} \\ &= \int_0^x \left(1 + \frac{1}{2}z^2 + \frac{1\cdot 3}{2\cdot 4}z^4 + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}z^6 + \cdots\right) dz,\end{aligned}$$

if $|z| < 1$, therefore

$$\sin^{-1} x = x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \cdots .$$

It is evident that this method of obtaining the expansion of $\sin^{-1} x$ is much less complicated than the direct application of Taylor's formula.

Example 3. In order to evaluate the integral

$$I = \int_0^h \frac{2a\,dx}{\sqrt{2g(h-x)(2ax-x^2)}},$$

express it as

$$\sqrt{\frac{a}{g}}\int_0^h \frac{dx}{\sqrt{hx - x^2}}\left(1 - \frac{x}{2a}\right)^{-\frac{1}{2}},$$

and then replace $\left(1 - \frac{x}{2a}\right)^{-\frac{1}{2}}$ by its expansion in powers of $\frac{x}{2a}$, giving

$$\sqrt{\frac{a}{g}}\int_0^h \left[1 + \frac{1}{2}\left(\frac{x}{2a}\right) + \frac{1 \cdot 3}{2 \cdot 4}\left(\frac{x}{2a}\right)^2 + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\left(\frac{x}{2a}\right)^3 + \cdots \right]\frac{dx}{\sqrt{hx - x^2}}.$$

If this integral is expressed as

$$\sqrt{\frac{a}{g}}\left[\int_0^h \frac{dx}{\sqrt{hx - x^2}} + \int_0^h \frac{1}{2}\left(\frac{x}{2a}\right)\frac{dx}{\sqrt{hx - x^2}} + \int_0^h \frac{1 \cdot 3}{2 \cdot 4}\left(\frac{x}{2a}\right)^2 \frac{dx}{\sqrt{hx - x^2}} + \cdots \right]$$

and each integral evaluated, there results

$$I = \pi\sqrt{\frac{a}{g}}\left[1 + \left(\frac{1}{2}\right)^2 \frac{h}{2a} + \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2\left(\frac{h}{2a}\right)^2 + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2\left(\frac{h}{2a}\right)^3 + \cdots \right].$$

This expression gives the period of the simple pendulum. By making the change of variable $x = h \sin^2 \varphi$, the integral reduces to

$$I = 2\sqrt{\frac{a}{g}}\int_0^{\frac{\pi}{2}} (1 - k^2 \sin^2 \varphi)^{-\frac{1}{2}}\, d\varphi,$$

where $k^2 = h/2a$.

This is the form used in the discussion of the simple pendulum given in Sec. 71. In this illustration, h denotes the height of the pendulum bob and a the length of the pendulum.

Example 4. The integral $\int_0^1 \frac{e^x - e^{-x}}{x}\, dx$ cannot be evaluated by the usual method for evaluating a definite integral, for the indefinite integral cannot be obtained. Moreover, the expansion for $\frac{e^x - e^{-x}}{x}$, if obtained directly with the aid of Maclaurin's formula, would lead to an extremely complicated expression for each derivative. The expansion is most easily obtained by using the separate expansions for e^x and e^{-x}. Thus,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots,$$

and

$$e^x - e^{-x} = 2\left(x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots\right).$$

Hence,

$$\int_0^1 \frac{e^x - e^{-x}}{x}\, dx = 2\left(1 + \frac{1}{3 \cdot 3!} + \frac{1}{5 \cdot 5!} + \cdots\right) = 2.1145.$$

Example 5. In order to evaluate the integral $\int_0^\pi e^{\sin x}\, dx$, recall that

$$e^u = 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \cdots,$$

so that

$$e^{\sin x} = 1 + \sin x + \frac{\sin^2 x}{2!} + \frac{\sin^3 x}{3!} + \cdots.$$

Then

$$\int_0^\pi e^{\sin x}\, dx = \int_0^\pi \left(1 + \sin x + \frac{\sin^2 x}{2!} + \frac{\sin^3 x}{3!} + \cdots\right) dx$$

$$= 2\int_0^{\frac{\pi}{2}} \left(1 + \sin x + \frac{\sin^2 x}{2!} + \frac{\sin^3 x}{3!} + \cdots\right) dx,$$

which can be evaluated with the aid of the Wallis formula

$$\int_0^{\frac{\pi}{2}} \sin^n x\, dx = \int_0^{\frac{\pi}{2}} \cos^n x\, dx = \frac{(n-1)(n-3)\cdots 2 \text{ or } 1}{n(n-2)\cdots 2 \text{ or } 1}\,\alpha,$$

where $\alpha = 1$, if n is odd, and $\alpha = \frac{\pi}{2}$, if n is even.

In order to justify the term-by-term integration, it is sufficient to show that the series in the integrand is uniformly convergent. That such is the case is obvious if one considers

$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots$$

as the Weierstrass M series.

PROBLEMS

1. Calculate cos 10°, and estimate the maximum error committed by neglecting terms after x^6.

2. Find the interval of convergence of the expansion of e^x in power series in x. Determine the number of terms necessary to compute $e^{1.1}$ accurate to four decimal places from this expansion, and compare the result with illustration 2, Sec. 12.

3. Compute sin 33°, correct to four decimal places.

4. Expand the integrand of $\int_0^x \frac{dx}{1+x^2}$ in power series in x, and integrate term by term. Compare the result with that of Prob. 1(d), Sec. 11.

5. Compute $\sqrt[5]{35} = 2(1 + 3/32)^{1/5}$, correct to five decimal places.

6. Develop the power series in x for $\sin^{-1} x$ and hence establish that

$$\frac{\pi}{6} = \frac{1}{2} + \frac{1}{2}\cdot\frac{1}{3}\left(\frac{1}{2}\right)^3 + \frac{1\cdot 3}{2\cdot 4}\cdot\frac{1}{5}\left(\frac{1}{2}\right)^5 + \cdots.$$

7. Show, by squaring and adding the power series for $\sin x$ and $\cos x$, that

$$\sin^2 x + \cos^2 x = 1.$$

8. Evaluate by using series expansions of the integrands

(a) $\int_0^1 \sin(x^2)\,dx$; (b) $\int_0^{1/5} \frac{\sin x\,dx}{\sqrt{1-x^2}}$;

(c) $\int_0^1 \frac{\sin x}{x}\,dx$; ($d$) $\int_0^x e^{-x^2}\,dx$;

(e) $\int_0^x \cos(x^2)\,dx$; (f) $\int_0^1 (2-\cos x)^{-1/2}\,dx$

$$= \int_0^1 \left[2 - \left(1 - 2\sin^2\frac{x}{2}\right)\right]^{-1/2} dx$$

$$= \int_0^1 \left(1 + 2\sin^2\frac{x}{2}\right)^{-1/2} dx;$$

(g) $\int_{0.9}^1 \frac{\log x}{1-x}\,dx = \int_0^{0.1} \frac{\log(1-z)}{z}\,dz$; ($h$) $\int_0^x \frac{\cos x\,dx}{\sqrt{x}}$;

(i) $\int_0^x e^{\tan x}\,dx$.*

9. Show, by multiplication of series, that

$$(1 + x + x^2 + \cdots)^2 = 1 + 2x + 3x^2 + \cdots$$
$$= (1-x)^{-2}.$$

10. Expand to terms in x^6

(a) $\sqrt{\cos x}$;

(b) $\frac{\sin x}{e^x - 1}$;

(c) $\frac{e^x}{1+e^x}$.

11. Determine the magnitude of α, if the error in the approximation $\sin\alpha \doteq \alpha$ is not to exceed 1 per cent.

Hint: $\frac{\alpha - \sin\alpha}{\alpha} = 0.01$ and $\sin\alpha = \alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \cdots.$

* See form 787, B. O. Peirce, A Short Table of Integrals.

14. Rectification of Ellipse. Elliptic Integrals. In spite of its importance and apparent simplicity, the problem of finding the length of an elliptical arc is not usually considered in elementary calculus. This is because the integral that arises is incapable of evaluation in terms of elementary functions. However, the evaluation can be effected by means of series expansion of the integrand function, as will be shown in this section.

Let the equation of the ellipse be

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \qquad a > b.$$

The length of arc from $(0, b)$ to (x_1, y_1) is given by the integral

$$s = \int_0^{x_1} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx. \tag{14-1}$$

Computing dy/dx and substituting its value in (14-1) gives

$$s = \int_0^{x_1} \sqrt{1 + \frac{b^2}{a^2}\frac{x^2}{a^2 - x^2}}\, dx = \int_0^{x_1} \sqrt{\frac{a^2 - \dfrac{a^2 - b^2}{a^2}x^2}{a^2 - x^2}}\, dx.$$

Recalling the fact that the numerical eccentricity of the ellipse is $k = \sqrt{a^2 - b^2}/a$, the integral given above can be written as

$$s = \int_0^{x_1} \sqrt{\frac{a^2 - k^2x^2}{a^2 - x^2}}\, dx, \tag{14-2}$$

where $k^2 < 1$.

Let $x = a \sin \theta$; then $dx = a \cos \theta\, d\theta$, and (14-2) becomes

$$s = a \int_0^{\varphi} \sqrt{1 - k^2 \sin^2 \theta}\, d\theta. \tag{14-3}$$

The series expansion of the integrand function is most easily obtained by writing it as $(1 - k^2 \sin^2 \theta)^{1/2}$ and expanding by use of the binomial theorem. Then (14-3) is replaced by

$$s = a \int_0^{\varphi} \left(1 - \frac{1}{2} k^2 \sin^2 \theta - \frac{1}{8} k^4 \sin^4 \theta - \cdots \right) d\theta,$$

and term-by-term integration* gives

* Term-by-term integration is valid here, for the series

$$1 + \frac{1}{2} k^2 + \cdots + \frac{1 \cdot 3 \cdots (2n-3)}{2 \cdot 4 \cdots 2n} k^{2n} + \cdots$$

serves as a Weierstrass M series.

$$(14\text{-}4)\qquad s = a\left[\varphi - \frac{1}{2}k^2\int_0^\varphi \sin^2\theta\,d\theta - \frac{1}{8}k^4\int_0^\varphi \sin^4\theta\,d\theta - \cdots \right.$$
$$\left. - \frac{1\cdot 3\cdot 5\cdots(2n-3)}{2\cdot 4\cdot 6\cdots 2n}k^{2n}\int_0^\varphi \sin^{2n}\theta\,d\theta - \cdots\right].$$

If (14-4) is used, it is possible to evaluate s for particular values of k and φ. However, the integral in (14-3) is so important that there are extensive tables* giving its value for many choices of k and φ. This integral for the value of $a = 1$ is called the elliptic integral of the second kind and is denoted by the symbol $E(k, \varphi)$. If $\varphi = \pi/2$, the integral is called the complete elliptic integral of the second kind, which is denoted by the symbol E.

The elliptic integral of the second kind having been defined, it seems desirable to mention the elliptic integral of the first kind, although the latter arises in considering the motion of a simple pendulum and will be discussed in more detail in Sec. 71. The elliptic integral of the first kind, $F(k, \varphi)$, has the form

$$(14\text{-}5)\qquad F(k, \varphi) = \int_0^\varphi \frac{d\theta}{\sqrt{1 - k^2\sin^2\theta}}.$$

The complete elliptic integral of the first kind, which arises when $\varphi = \pi/2$, is denoted by the symbol K. Values of $F(k, \varphi)$ and of K are also tabulated, but the evaluation can be obtained from (14-5) by means of series expansion of the integrand. Thus, one has the expansion

$$(14\text{-}6)\qquad F(k, \varphi) = \varphi + \frac{1}{2}k^2\int_0^\varphi \sin^2\theta\,d\theta + \frac{3}{8}k^4\int_0^\varphi \sin^4\theta\,d\theta$$
$$+ \cdots + \frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n}k^{2n}\int_0^\varphi \sin^{2n}\theta\,d\theta + \cdots.$$

15. Discussion of Elliptic Integrals. The elliptic integral of the first kind is a function defined by the integral

$$(15\text{-}1)\qquad F(k, \varphi) \equiv \int_0^\varphi \frac{d\theta}{\sqrt{1 - k^2\sin^2\theta}},\qquad k^2 < 1.$$

* See the brief table in B. O. Peirce, A Short Table of Integrals, pp. 121–123.

If $\sin\theta$ is replaced by z, (15-1) becomes

$$\bar{F}(k, x) = \int_0^x \frac{dz}{\sqrt{(1-z^2)(1-k^2z^2)}}, \qquad k^2 < 1. \tag{15-2}$$

This is an alternative form of the elliptic integral of the first kind.

Similarly, the same change of variable transforms the integral of the second kind

$$E(k, \varphi) = \int_0^\varphi \sqrt{1 - k^2 \sin^2\theta}\, d\theta, \qquad k^2 < 1, \tag{15-3}$$

into

$$\bar{E}(k, x) = \int_0^x \sqrt{\frac{1-k^2z^2}{1-z^2}}\, dz, \qquad k^2 < 1. \tag{15-4}$$

It will be recalled that any integral of the type

$$\int R(x, \sqrt{ax^2 + bx + c})\, dx,$$

where R is a rational function of the variables x and

$$\sqrt{ax^2 + bx + c},$$

is integrable in terms of the elementary functions, *i.e.*, power, trigonometric, and logarithmic functions. It can be shown that the integration of integrals of the type

$$\int R(x, \sqrt{ax^3 + bx^2 + cx + d})\, dx \tag{15-5}$$

and

$$\int R(x, \sqrt{ax^4 + bx^3 + cx^2 + dx + e})\, dx \tag{15-6}$$

requires, in general, the introduction of new functions obtained from the elliptic integrals.

The evaluation of (15-5) and (15-6) can be reduced to the evaluation of integrals of the elementary types and the following new types:

a. Elliptic integral of the first kind:

$$\bar{F}(k, x) \equiv \int_0^x \frac{dz}{\sqrt{(1-z^2)(1-k^2z^2)}}, \qquad \text{or}$$

$$F(k, \varphi) \equiv \int_0^\varphi \frac{d\theta}{\sqrt{1 - k^2\sin^2\theta}}.$$

b. Elliptic integral of the second kind:

$$\bar{E}(k, x) \equiv \int_0^x \sqrt{\frac{1-k^2z^2}{1-z^2}}\, dz, \qquad \text{or}$$

$$E(k, \varphi) \equiv \int_0^\varphi \sqrt{1 - k^2\sin^2\theta}\, d\theta.$$

c. Elliptic integral of the third kind:

$$\bar{\Pi}(n, k, x) \equiv \int_0^x \frac{dz}{(1 + nz^2)\sqrt{(1 - z^2)(1 - k^2z^2)}},$$

or

$$\Pi(n, k, \varphi) \equiv \int_0^\varphi \frac{d\theta}{(1 + n \sin^2 \theta)\sqrt{1 - k^2 \sin^2 \theta}}.$$

The problem of reducing the integrals of expressions involving square roots of cubics and quartics to normal forms is not difficult, but it is tedious and will be omitted here.* Integrals involving

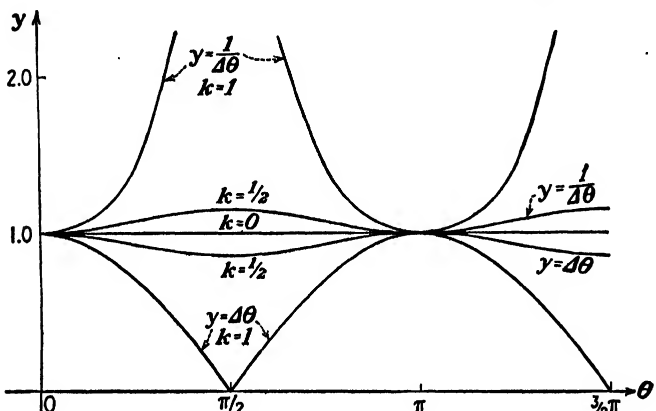

FIG. 4.

square roots of polynomials of degree higher than the fourth lead, in general, to more complicated functions, the so-called hyper-elliptic functions.

The graphs of the integrands of the integrals of the first and second kinds are of some interest (see Fig. 4). For $k = 0$,

$$\Delta\theta \equiv \sqrt{1 - k^2 \sin^2 \theta} \qquad \text{and} \qquad \frac{1}{\Delta\theta} \equiv \frac{1}{\sqrt{1 - k^2 \sin^2 \theta}}$$

both become equal to 1, and the corresponding integrals are both equal to φ. For $0 < k < 1$, the curve $y = 1/\Delta\theta$ lies entirely above the line $y = 1$ and the curve $y = \Delta\theta$ lies entirely below it. As φ increases,

* For a detailed account see Goursat-Hedrick, Mathematical Analysis, vol. 1, p. 226. A monograph, Elliptic Functions by H. Hancock, may also be consulted.

$F(k, \varphi)$ and $E(k, \varphi)$ increase continuously, F being always the larger. As k increases, φ being fixed, the value of $F(k, \varphi)$ increases and that of $E(k, \varphi)$ decreases. Also $F(k, \pi) = 2K$ and $E(k, \pi) = 2E$, for the curves are symmetrical about $\theta = \pi/2$. If $\pi/2 < \varphi < \pi$, it is obvious from the figure that

(15-7)
$$F(k, \varphi) = 2K - F(k, \pi - \varphi),$$
$$E(k, \varphi) = 2E - E(k, \pi - \varphi).$$

Moreover,

(15-8)
$$F(k, m\pi + \varphi) = 2mK + F(k, \varphi),$$
$$E(k, m\pi + \varphi) = 2mE + E(k, \varphi),$$

where m is an integer.

Since the values of K and E, and of $F(k, \varphi)$ and $E(k, \varphi)$ for $\varphi \leq \pi/2$, are tabulated, the relations (15-7) and (15-8) permit the evaluation of $F(k, \varphi)$ and $E(k, \varphi)$ for all values of φ.

The discussion* above was restricted to values of $k^2 < 1$. If $k^2 = 1$, $y = \Delta\theta$ becomes $y = |\cos\theta|$ and $y = 1/\Delta\theta$ becomes $y = |\sec\theta|$.

Consider

(15-9)
$$u = \int_0^x \frac{dz}{\sqrt{(1 - z^2)(1 - k^2z^2)}} = \int_0^\varphi \frac{d\theta}{\sqrt{1 - k^2\sin^2\theta}},$$

where $x = \sin\varphi$.

For a fixed value of k, (15-9) defines $u = \bar{F}(x)$ or $u = F(\varphi)$. The function resulting from the solution of (15-9) for φ in terms of u is called the amplitude of u and is denoted by am $(u, \text{mod } k)$, or more simply by $\varphi = \text{am } u$. It will be assumed that the equation $u = F(\varphi)$ can be solved for φ. Since $\varphi = \text{am } u$,

$$x = \sin\varphi \equiv \sin \text{am } u \equiv \text{sn } u.$$

Moreover,

$$\cos\varphi \equiv \sqrt{1 - x^2} \equiv \sqrt{1 - \text{sn}^2 u} \equiv \text{cn } u.$$

Finally,

$$\Delta\varphi \equiv \sqrt{1 - k^2x^2} \equiv \text{dn } u.$$

The functions sn u, cn u, and dn u are called the *elliptic functions*.

From the definitions, it is obvious that

$$\text{am }(0) = 0, \quad \text{sn }(0) = 0, \quad \text{cn }(0) = 1, \quad \text{dn }(0) = 1;$$
$$\text{am }(-u) = -\text{am } u, \text{ sn }(-u) = -\text{sn } u, \text{ cn }(-u) = \text{cn } u, \text{ dn }(-u) = \text{dn } u.$$

The elliptic functions are periodic functions and in some respects resemble the trigonometric functions. There exists a complete set of

* See Prob. 1, at the end of this section.

formulas connecting the elliptic functions analogous to the set for the trigonometric functions.*

An interesting application of elliptic integrals to electrical problems is found in the calculation of the magnetic flux density in the plane of a circular loop of radius a carrying a steady current I.

Upon applying the law of Biot and Savart† to a circular loop of radius a, the flux density B at any point P in the plane of the wire is given by

$$B = \frac{I}{4\pi}\int_C \frac{\sin(r, s)\, ds}{r^2}, \tag{15-10}$$

where C is the circumference of the loop, r is the radius vector from P to an element of arc ds, and (r, s) is the angle between r and this element (Fig. 5).

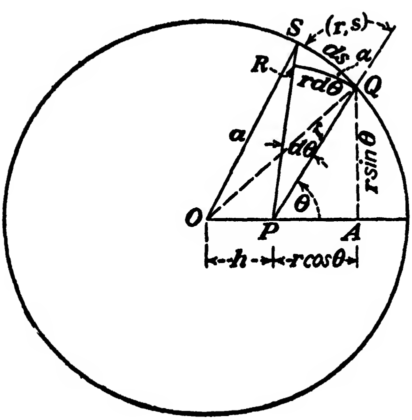

Fig. 5.

If the point P is at the center of the loop, then $(r, s) = 90°$, $r = a$, and the integral is easily evaluated to give

$$B = \frac{I}{4\pi}\frac{2\pi a}{a^2} = \frac{I}{2a},$$

a familiar result.

If, however, the point P is not at the center, the evaluation of the integral is not so easy. Consider the triangle RQS, where the side $RQ = r\,d\theta$ makes an angle α with ds. It is clear that $ds\cos\alpha = r\,d\theta$; and, since $\alpha = 90° - (r, s)$, it follows that

$$\cos\alpha = \sin(r, s).$$

Hence,

$$ds = \frac{r\,d\theta}{\sin(r, s)}.$$

* See Appel, P., and E. Lacour, Fonctions elliptiques; Peirce, B. O., A Short Table of Integrals; Greenhill, A. G., The Application of Elliptic Functions.

† This formula is known to engineers as Ampere's formula. See, for example, E. Bennett, Introductory Electrodynamics for Engineers. The system of units used here is the "rational" system of units used in M. Mason and W. Weaver, Electromagnetic Field.

The substitution of this value in (15-10) yields

$$(15\text{-}11)\qquad B = \frac{I}{4\pi}\int_0^{2\pi}\frac{d\theta}{r}$$

for the magnetic flux density at P.

Now, from triangle OQA, it is evident that

$$\sqrt{a^2 - (r\sin\theta)^2} = r\cos\theta + h,$$

which, after squaring both sides and simplifying, becomes

$$r^2 + 2rh\cos\theta + (h^2 - a^2) = 0.$$

Solving for r gives

$$r = -h\cos\theta \pm \sqrt{h^2\cos^2\theta + a^2 - h^2};$$

and, since r is always positive, the radical must be taken with the positive sign. Substituting this value of r in (15-11) gives

$$B = \frac{I}{4\pi}\int_0^{2\pi}\frac{d\theta}{-h\cos\theta + \sqrt{h^2\cos^2\theta + a^2 - h^2}},$$

or, upon rationalization of the denominator,

$$\begin{aligned} B &= \frac{I}{4\pi}\int_0^{2\pi}\frac{-h\cos\theta - \sqrt{h^2\cos^2\theta + a^2 - h^2}}{h^2 - a^2}\,d\theta \\ &= \frac{I}{4\pi(a^2 - h^2)}\left(\int_0^{2\pi} h\cos\theta\,d\theta + \int_0^{2\pi}\sqrt{a^2 - h^2\sin^2\theta}\,d\theta\right). \end{aligned}$$

The first of these integrals is zero, and the second is an elliptic integral of the second kind, so that

$$\begin{aligned} B &= \frac{Ia}{4\pi(a^2 - h^2)}\int_0^{2\pi}\sqrt{1 - \frac{h^2}{a^2}\sin^2\theta}\,d\theta \\ &= \frac{Ia}{\pi(a^2 - h^2)}\int_0^{\frac{\pi}{2}}\sqrt{1 - k^2\sin^2\theta}\,d\theta, \end{aligned}$$

where $k = h/a$. This integral can be evaluated for any k with the aid of the tables of elliptic integrals.

PROBLEMS

1. Prove that

$$\int_0^{\varphi}\frac{d\varphi}{\sqrt{1 - l^2\sin^2\varphi}} = \frac{1}{l}\int_0^{\psi}\frac{d\alpha}{\sqrt{1 - l^{-2}\sin^2\alpha}},\qquad l > 1.$$

Hint: Change the variable by setting $l^2\sin^2\varphi = \sin^2\alpha$.

2. Plot, with the aid of Peirce's tables, $F(k, \varphi)$, where $k = \sin \alpha$, using α as abscissa and $F(k, \varphi)$ as ordinate. Draw 10 curves on the same sheet of rectangular coordinate paper for $\varphi = 0$, $\varphi = 10$, $\varphi = 20$, $\varphi = 30$, $\varphi = 40$, $\varphi = 50$, $\cdots$, $\varphi = 90$.

3. Plot four curves representing $F(k, \varphi)$ on the same sheet of rectangular coordinate paper. Use φ as abscissa and the values of k as 0, ½, $\sqrt{3}/2$, and 1.

4. Plot the integrand of $\int_0^\varphi \frac{d\varphi}{\sqrt{1 - k^2 \sin^2 \varphi}}$ for the values of $k = 0$, ½, and 1. Use φ as abscissa. The areas under the curves give the values of the elliptic integrals.

5. Compute the value of $F(0, \pi/2)$.

6. The major and minor axes of an elliptical arch are 200 ft. and 50 ft., respectively. Find the length of the arch. Compute the length of the arch between the points where $x = 0$ and $x = 25$. Use Peirce's tables.

7. Plot with the aid of Peirce's tables $E(k, \varphi)$, where $k = \sin \alpha$. Use α's as abscissas and $E(k, \varphi)$ as ordinates. Draw 10 curves on the same sheet of rectangular coordinate paper for $\varphi = 0, 10, 20, \cdots, 90$.

8. Plot on a sheet of rectangular coordinate paper the four curves representing $E(k, \varphi)$. Use φ as abscissa. The four curves are for $k = 0$, ½, $\sqrt{3}/2$, and 1.

9. Plot the integrand of $\int_0^\varphi \sqrt{1 - k^2 \sin^2 \varphi}\, d\varphi$ for the values of $k = 0$, ½, and 1. Use φ as abscissa. Compare the result with that of Prob. 4. What can be said about the relative magnitudes of $F(k, \varphi)$ and $E(k, \varphi)$?

10. Show that $\int_0^\varphi \frac{d\varphi}{\sqrt{1 + k^2 \sin^2 \varphi}}$ is an elliptic integral of the first kind.

Hint: Change the variable by setting $\sin \varphi = \frac{1}{k} \tan x$.

11. Show that

$$\int_0^{\frac{\pi}{2}} \frac{dx}{\sqrt{\sin x}} = \int_0^{\frac{\pi}{2}} \frac{dx}{\sqrt{\cos x}} = \sqrt{2} \int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{1 - \frac{1}{2} \sin^2 \varphi}}.$$

Hint: Set $\sqrt{\cos x} = \cos \varphi$.

Note that the integral is improper but that it is easy to show its convergence.

12. Show that

$$\int_0^{\frac{\pi}{2}} \frac{\sin^2 \theta \, d\theta}{\sqrt{1 - k^2 \sin^2 \theta}} = \frac{1}{k^2} (K - E).$$

Hint: $\sin^2\theta = \frac{1}{k^2} - \frac{1}{k^2}(1 - k^2\sin^2\theta)$.

13. Show that

$$K \equiv \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1 - k^2\sin^2\theta}}$$
$$= \frac{\pi}{2}\left[1 + \left(\frac{1}{2}\right)^2 k^2 + \left(\frac{1\cdot 3}{2\cdot 4}\right)^2 k^4 + \left(\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right)^2 k^6 + \cdots\right] \text{ if } k^2 < 1.$$

14. Find the length of one arch of the sine curve.

15. Find the length of the portion of $y = \sin x$ lying between $x = 1$ and $x = 2$.

16. Given:

$$F\left(\frac{1}{2}, \varphi\right) \equiv \int_0^{\varphi} \frac{d\theta}{\sqrt{1 - \frac{1}{4}\sin^2\theta}}.$$

Find K and $\operatorname{sn} \frac{2}{3}K$.

17. Show that $\int \frac{d\theta}{\sqrt{a - \cos\theta}}$, where $a > 1$, is an elliptic integral.

18. Show that the length of arc of an ellipse of semiaxes a and b is given by

$$s = 4a\int_0^{\frac{\pi}{2}} \sqrt{1 - e^2\sin^2\theta}\, d\theta$$
$$= 2\pi a\left(1 - \frac{e^2}{4} - \frac{3}{64}e^4 - \cdots\right), \text{ where } e \text{ is the eccentricity.}$$

16. Approximate Formulas in Applied Mathematics. It is frequently necessary to introduce approximations in order to make readily usable the results of mathematical investigations. For example, an engineer seldom finds it necessary to use the exact formula for the curvature of a curve whose equation is $y = f(x)$, namely,

$$K = \frac{\frac{d^2y}{dx^2}}{\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{\frac{3}{2}}}, \tag{16-1}$$

since in most applications the slope dy/dx is small enough to permit the use of the approximate formula

$$K \doteq \frac{d^2y}{dx^2}. \tag{16-2}$$

Many such approximations are obtained by using the first few terms of the Taylor's series expansion in place of the function

itself. Thus, the formula (16-2) is obtained from (16-1) by neglecting all except the first term in the expansion of $[1 + (dy/dx)^2]^{-3/2}$ in powers of dy/dx.

1. *Small Errors.* The values of physical quantities determined by experiment are subject to errors due to inaccuracies arising in the measurements of the quantities involved. It is often necessary to know the size of such errors.

Let a capillary tube contain a column of mercury. The radius R of the tube can be determined by measuring the length L and the weight W of the column of mercury. Let L be measured in centimeters and W in grams. Since the density of mercury is $\rho = 13.6$,

$$R = \sqrt{\frac{W}{\pi \rho L}} = 0.153 \sqrt{\frac{W}{L}}.$$

The principal error arises in the measurement of L. Let L be the true value, and let $L' = L + \epsilon$ be the observed value. Then, if R is the true value of the radius, let $R' = R + \eta$ be the computed value. The error in measuring W is negligible because of the high accuracy of the balance. It follows that

$$R = 0.153 \sqrt{\frac{W}{L}} \qquad \text{and} \qquad R' = 0.153 \sqrt{\frac{W}{L'}}$$

or

$$R + \eta = 0.153 \sqrt{\frac{W}{L + \epsilon}}.$$

Therefore,

$$\begin{aligned} \eta &= 0.153 \left(\sqrt{\frac{W}{L + \epsilon}} - \sqrt{\frac{W}{L}} \right) \\ &= 0.153 \sqrt{\frac{W}{L}} \left[\left(1 + \frac{\epsilon}{L} \right)^{-1/2} - 1 \right] \\ &= R \left(-\frac{\epsilon}{2L} + \frac{3}{8} \frac{\epsilon^2}{L^2} - \cdots \right). \end{aligned}$$

Since ϵ is small compared with L, it follows that η is approximately given by $-\frac{1}{2} R \frac{\epsilon}{L}$. Clearly, ϵ can be either positive or negative.

2. *Crank and Connecting Rod.* If one end of a straight line PQ (see Fig. 6) is required to move on a circle, while the other

end moves on a straight line which passes through the center of the circle, the resulting motion is called connecting-rod motion. This kind of motion arises in a steam engine in which one end of the connecting rod is attached to the crank PB and therefore moves in a circle whose radius is the length of the crank, while the other end is attached to the crosshead and moves along a straight line.

Let r be the length of the crank, l the length of the connecting rod, and s the displacement of the crosshead from the position A,

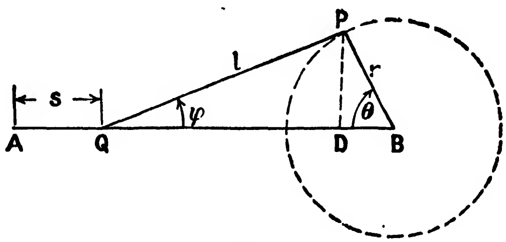

FIG. 6.

in which the connecting rod and crank lie in a straight line. Then,

$$AB = l + r,$$

and

$$AB = AQ + QD + DB = s + l \cos \varphi + r \cos \theta.$$

Moreover,

$$PD = l \sin \varphi = r \sin \theta,$$

so that

$$\sin \varphi = \frac{r}{l} \sin \theta$$

and

$$\cos \varphi = \sqrt{1 - \frac{r^2}{l^2} \sin^2 \theta}.$$

Therefore,

$$s + l\left(1 - \frac{r^2}{l^2} \sin^2 \theta\right)^{1/2} + r \cos \theta = l + r$$

or

$$s = l\left[1 - \left(1 - \frac{r^2}{l^2} \sin^2 \theta\right)^{1/2}\right] + r(1 - \cos \theta).$$

If

$$\left(1 - \frac{r^2}{l^2} \sin^2 \theta\right)^{1/2}$$

be replaced by its expansion, it follows that

$$s = l\left[\frac{1}{2}\frac{r^2}{l^2}\sin^2\theta + \frac{1}{8}\left(\frac{r^2}{l^2}\right)^2\sin^4\theta + \cdots\right] + r(1 - \cos\theta)$$
$$= \left(\frac{r^2}{2l}\sin^2\theta + \frac{r^4}{8l^3}\sin^4\theta + \cdots\right) + r(1 - \cos\theta).$$

If r is small compared with l, the displacement of the crosshead is given approximately by $r(1 - \cos\theta)$.

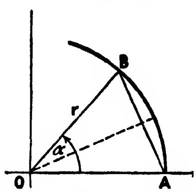

FIG. 7.

3. *Surveying.* In railroad surveying, it is frequently useful to know the amount of difference between the length of a circular arc and the length of its corresponding chord.

Let r be the radius of curvature of the arc AB (Fig. 7), and let α be the angle intercepted by the arc. Then, if s is the length of the arc AB and c is the length of chord AB, $s = r\alpha$ and $c = 2r\sin\frac{\alpha}{2}$. Since

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!}\cos\xi,$$

where $0 < \xi < x$, the error in using only the first two terms of the expansion is certainly less than $\left|\frac{x^5}{5!}\right|$. Then,

$$c = 2r\sin\frac{\alpha}{2} = 2r\left(\frac{\alpha}{2} - \frac{\alpha^3}{8\cdot 6}\right)$$

with an error less than

$$2r\left(\frac{\alpha^5}{32\cdot 120}\right) = \frac{r\alpha^5}{1920}.$$

Therefore,

$$s - c = r\alpha - r\alpha + \frac{\alpha^3 r}{24} = \frac{\alpha^3 r}{24}$$

with an error that is less than $r\alpha^5/1920$.

4. *Vertical Motion under Earth's Attraction.* Let it be required to determine the velocity of a body of mass m that is falling from a height s_0 above the center of the earth and is subject to the earth's attraction alone.

Let F be the attraction on the earth's surface and F' be the attraction at a distance h from the surface (Fig. 8). Then

$$F = \frac{kmm'}{r^2} \qquad \text{and} \qquad F' = \frac{kmm'}{(r+h)^2},$$

where m' is the mass of the earth, k is the gravitational constant, and r is the radius of the earth. Hence,

$$\frac{F}{F'} = \frac{(r+h)^2}{r^2}.$$

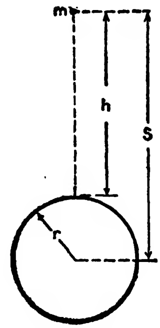

FIG. 8.

Also, let g be the acceleration at the surface of the earth and g' be the acceleration at a distance h above the surface, so that $F = mg$ and $F' = mg'$. It follows that

$$\frac{F}{F'} = \frac{g}{g'} = \frac{(r+h)^2}{r^2} = \frac{s^2}{r^2},$$

and, therefore,

$$g' = \frac{gr^2}{s^2}.$$

But

$$g' = -\frac{d^2s}{dt^2},$$

so that

$$\frac{d^2s}{dt^2} = -\frac{gr^2}{s^2}.$$

This equation can be solved for $v = ds/dt$ by the following device: Multiplying both members by $2\,ds/dt$ and integrating give

$$\left(\frac{ds}{dt}\right)^2 = \frac{2gr^2}{s} + C,$$

where C is the constant of integration. If the initial velocity $(ds/dt)_{s=s_0}$ is zero, then $C = -2gr^2/s_0$ and hence

$$\left(\frac{ds}{dt}\right)^2 = 2gr^2\left(\frac{1}{s} - \frac{1}{s_0}\right).$$

But $s = r + h$ and $ds/dt = v$, so that the equation becomes

$$v^2 = 2gr^2\left(\frac{1}{r+h} - \frac{1}{s_0}\right).$$

This formula can be used to calculate the terminal velocity (*i.e.*, the velocity at the earth's surface) when the body is released

from any height. Thus, setting $h = 0$ gives

$$v^2 = 2gr^2\left(\frac{1}{r} - \frac{1}{s_0}\right). \tag{16-3}$$

Upon denoting by h_0 the initial height above the earth's surface, so that $s_0 = r + h_0$, (16-3) can be written as

$$v^2 = 2gr^2\left(\frac{1}{r} - \frac{1}{r + h_0}\right),$$

or

$$v^2 = 2gr\left(1 - \frac{r}{r + h_0}\right). \tag{16-4}$$

Now $\dfrac{r}{r + h_0} = \left(1 + \dfrac{h_0}{r}\right)^{-1}$; and if $\dfrac{h_0}{r} < 1$, then series expansion is permissible, so that

$$\frac{r}{r + h_0} = 1 - \frac{h_0}{r} + \left(\frac{h_0}{r}\right)^2 - \cdots.$$

Hence, if $h_0/r < 1$, (16-4) can be replaced by

$$v^2 = 2gr\left[\frac{h_0}{r} - \left(\frac{h_0}{r}\right)^2 + \left(\frac{h_0}{r}\right)^3 - \cdots\right].$$

Moreover, if h_0 is very small compared with r, then the powers of h_0/r higher than the first can be neglected* and

$$v^2 = 2gh_0,$$

which is the familiar formula for the terminal velocity of a body falling freely from a height h_0 that is not too great.

It follows from (16-3) that the square of the terminal velocity will be less than $2gr^2(1/r) = 2gr$. Moreover, for large values of s_0 the terminal velocity will be very close to $\sqrt{2gr}$. Accordingly, if a body falls from a very great distance it would attain a terminal velocity (air resistance being neglected) of approximately

$$\sqrt{2gr} = 6.95 \text{ miles per second.}$$

The results stated in the last paragraph may receive a different interpretation. Suppose a body were projected outward from the earth's surface with a velocity of more than $\sqrt{2gr} = 6.95$ miles per second. The previous discussion shows that, if air

* Since the series is alternating, the error will be less than $2gr(h_0/r)^2$.

resistance is neglected, the body would travel an infinite distance. This velocity is called the critical velocity or the velocity of escape.

It may be recalled that the earth's rotation exerts a centrifugal force on a particle which is falling toward the earth and that this force diminishes the effect of the force due to the earth's attraction. For a particle of mass m on the surface of the earth at the equator, this centrifugal force is

$$\frac{mv^2}{r} = \frac{m\omega^2r^2}{r} = m\omega^2r = \frac{mg}{289} \text{ dynes,}$$

where $\omega = 0.00007292$ radian per second is the angular velocity of the earth, $r = 6{,}370{,}284$ m., and $g = 980$ cm. per second per second. At a distance s from the center of the earth, this force is

$$m\omega^2s = \frac{mgs}{289r}.$$

But the earth's attraction at this distance is $F = mg'$. Since $g' = gr^2/s^2$,

$$F = \frac{mgr^2}{s^2}.$$

If the particle is to be in equilibrium,

$$\frac{mgs}{289r} = \frac{mgr^2}{s^2},$$

so that

$$s^3 = 289r^3 \qquad \text{or} \qquad s = 6.6r = 26{,}000 \text{ miles approx.}$$

Thus, if all other forces are neglected, a particle would be in equilibrium at approximately 22,000 miles above the earth's surface. This gives a very rough approximation to the extent of the earth's atmosphere. The actual thickness of the atmospheric layer is supposed to be considerably smaller.

PROBLEMS

1. The mass of the moon is nearly one-eighty-first that of the earth, and its radius is approximately three-elevenths that of the earth. Determine the velocity of escape for a body projected from the moon. Acceleration of gravity on the surface of the moon is one-sixth that on the surface of the earth.

2. Show that the time required for a body to reach the surface of the earth in Illustration 4, Sec. 16, is

$$t = \frac{\sqrt{s_0}}{8r}\left(\sqrt{s_0 s - s^2} + \frac{s_0}{2}\cos^{-1}\frac{2s - s_0}{s_0}\right).$$

Hint:

$$\frac{ds}{dt} = -\sqrt{2gr^2}\left(\frac{1}{s} - \frac{1}{s_0}\right)^{1/2}.$$

3. If the earth is considered as a homogeneous sphere at rest, then the force of attraction on a particle within the sphere can be shown to be proportional to the distance of the particle from the center. Let a hole be bored through the center of the earth, the air exhausted, and a stone released from rest at the surface of the earth. Show that the velocity of the stone at the center of the earth is about 5 miles per second.

Hint:

$$m\frac{d^2s}{dt^2} = -\frac{mg}{r}s,$$

where s is the distance of the stone from the center of the earth and r is the radius of the earth.

CHAPTER II

FOURIER SERIES

17. Preliminary Remarks. It is frequently necessary to find the equation of a curve that passes through a certain number of given points lying in the xy-plane. This can be accomplished in an infinite number of ways. Thus, if there are three given points, the coefficients of

$$y = a_0 + a_1x + a_2x^2 \tag{17-1}$$

can be chosen so that the resulting parabola will pass through the three given points. This is accomplished by solving the three linear equations, in a_0, a_1, and a_2, that arise when the coordinates of the given points are substituted in (17-1). If there are four given points, it is impossible, in general, to determine a_0, a_1, and a_2 so that the parabola (17-1) will pass through all four points, since the four linear equations in a_0, a_1, and a_2 will be, in general, incompatible. However, it will be possible to determine the coefficients of

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 \tag{17-2}$$

so that the curve defined by (17-2) passes through all four points.

The determination of the equation of a curve passing through a set of given points is not unique. Thus, if four points are given, it is possible to determine a curve whose equation is (17-2) which passes through them. But it is also possible to determine the coefficients of

$$y = b_0 + b_1x + b_2 \sin x + b_3x^5$$

so that the curve defined by this equation will pass through the given points. Obviously this curve will not coincide with that defined by (17-2). The type of curve can be varied at will, but the number of coefficients to be determined must equal the number of given points.

If a curve is defined by the equation $y = f(x)$, it is possible, as indicated above, to make the curve whose equation is

$$\text{(17-3)} \quad y = a_0 + a_1 \cos x + a_2 \cos 2x + \cdots + a_n \cos nx + b_1 \sin x + b_2 \sin 2x + \cdots + b_n \sin nx$$

pass through any $2n + 1$ points of $y = f(x)$ in the interval from $x = 0$ to $x = 2\pi$. The question arises as to whether it is possible to make the curve $y = f(x)$ coincide with that defined by (17-3) at all points of the interval $(0, 2\pi)$ by increasing indefinitely the number of terms in (17-3). It is already known that it is possible, under rather severe restrictions, to represent $f(x)$ by an infinite series in powers of x. This was accomplished with the aid of Taylor's series. The analogous problem of representing $f(x)$ by an infinite trigonometric series was developed by Fourier and will be discussed in the succeeding sections.

Whereas the representation of a function $f(x)$ in Taylor's series demands that $f(x)$ possess derivatives of all orders, the development in a trigonometric series

$$\text{(17-4)} \qquad \sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx)$$

is possible for a much larger group of functions. In fact, many periodic* functions having a finite number of ordinary discontinuities can be represented by infinite series of trigonometric functions. The term ordinary (or finite) discontinuity is used to describe the situation that arises when the function $f(x)$ suffers a finite jump at some point $x = x_0$ (see Fig. 9). Analytically, this means that the two limiting values of $f(x)$, as x approaches x_0 from the right-hand and the left-hand sides, exist but are unequal; *i.e.*,

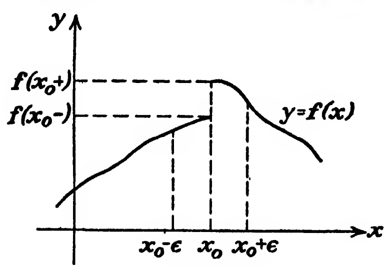

FIG. 9.

$$\lim_{\epsilon \to 0} f(x_0 + \epsilon) \neq \lim_{\epsilon \to 0} f(x_0 - \epsilon).$$

In order to economize on space, these right-hand and left-hand limits are written as $f(x_0+)$ and $f(x_0-)$, respectively, so that the foregoing inequality can be written as

$$f(x_0+) \neq f(x_0-).$$

* A function is said to be periodic of period a if $f(x) = f(x + a)$.

Inasmuch as each term of (17-4) is a periodic function of period 2π, it is necessary to restrict the discussion of the representation of functions by series of the type (17-4) to those functions which have period 2π. Or, what amounts to the same thing, the problem of representing a non-periodic function can be restricted to some interval of width 2π, and the function defined outside this interval so that it is periodic. For the present, it will be assumed that the interval in which $f(x)$ is considered is the interval $(-\pi, \pi)$ and that outside this interval the function is defined by the equation $f(x + 2\pi) = f(x)$. Of course, any interval $(a, a + 2\pi)$ would do equally well.

The theory of Fourier series is one of the most beautiful developments of analysis, and it serves as an indispensable instrument in the treatment of nearly every physical problem. Solutions of such important problems as sound vibration, propagation of electric currents and wireless waves, heat conduction, and mechanical vibrations give but a mere indication of its value.

The following section contains the celebrated theorem giving conditions on $f(x)$ that are sufficient to permit its representation by a Fourier series, and also a derivation* of the formulas for the Fourier coefficients (that is, the coefficients in the trigonometric series).

18. Dirichlet Conditions. Derivation of Fourier Coefficients. THEOREM. *Let $f(x)$ be a function defined arbitrarily in the interval $-\pi \leq x \leq \pi$, and outside this interval defined by the equation $f(x + 2\pi) = f(x)$. If $f(x)$ has a finite number of points of ordinary discontinuity and a finite number of maxima and minima in the interval $-\pi \leq x \leq \pi$, then it can be represented by the series*

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx),$$

with

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos kt \, dt,$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin kt \, dt, \qquad (k = 0, 1, 2, \cdots),$$

* For a more extended treatment, see I. S. Sokolnikoff, Advanced Calculus, Chap. XI.

*which converges at every point $x = x_0$ of the interval to the value**

$$\frac{f(x_0+) + f(x_0-)}{2}.$$

The restrictions imposed upon the function $f(x)$ in this theorem are known as the *Dirichlet conditions.*

The following demonstration that the Fourier coefficients a_k and b_k have the form given in the theorem assumes that the Fourier series development of the function $f(x)$, namely,

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \tag{18-1}$$

can be integrated term by term in the interval $(-\pi, \pi)$. The proof that if a function $f(x)$ satisfies the Dirichlet conditions then its Fourier series expansion actually converges to $f(x)$ is too involved to be given here.†

In order to determine a_0, multiply (18-1) by dx and integrate term by term from $-\pi$ to π. Since

$$\int_{-\pi}^{\pi} \cos nx\, dx = \int_{-\pi}^{\pi} \sin nx\, dx = 0 \text{ for } n = 1, 2, \cdots$$

and, hence,

$$\int_{-\pi}^{\pi} f(x)\, dx = a_0\pi,$$

there results

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, dx. \tag{18-2}$$

The coefficient a_n of the general cosine term can be obtained by multiplying both members of (18-1) by $\cos nx\, dx$ and performing term-by-term integration from $-\pi$ to π. Since, for all integral values of m and n,

$$\int_{-\pi}^{\pi} \sin mx \cos nx\, dx = 0$$

and

$$\int_{-\pi}^{\pi} \cos mx \cos nx\, dx = 0, \qquad \text{for } m \neq n,$$

* If $f(x)$ is continuous at the point $x = x_0$, then $f(x_0+) = f(x_0-) = f(x_0)$, so that at all points of continuity the series converges to $f(x)$. At the points of ordinary discontinuity, it converges to the arithmetic mean of the values of the right- and left-hand limits.

† KNOPP, K., Theory and Application of Infinite Series, p. 356; CARSLAW, H. S., Fourier's Series and Integrals, p. 207.

there results

$$\int_{-\pi}^{\pi} f(x) \cos nx\, dx = a_n \int_{-\pi}^{\pi} \cos^2 nx\, dx = a_n\pi.$$

Therefore,

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx. \tag{18-3}$$

It should be observed that (18-3) becomes (18-2) for $n = 0$.

Similarly, by multiplying (18-1) by $\sin nx\, dx$ and performing term-by-term integration from $-\pi$ to π, one obtains

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx\, dx. \tag{18-4}$$

It can be shown that if the values of the function are not equal at the end points of the interval $(-\pi, \pi)$, that is, if $f(-\pi) \neq f(\pi)$, then at these end points the Fourier series expansion for $f(x)$ converges to $\frac{1}{2}[f(-\pi) + f(\pi)]$.

The student will convince himself that, if the function $f(x)$ is defined in the interval from 0 to 2π, then the coefficients a_n and b_n in (18-1) are given by the formulas

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx\, dx \quad \text{and} \quad b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx\, dx.$$

19. Expansion of Functions in Fourier Series. This section contains some illustrative examples of expansion of functions, satisfying the Dirichlet conditions in the interval $(-\pi, \pi)$, in the series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{19-1}$$

where the coefficients a_n and b_n are given by the formulas

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx \tag{19-2}$$

and

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx\, dx. \tag{19-3}$$

Illustrative Example 1. Expand $f(x) = x$ in Fourier series in the interval $-\pi \leq x \leq \pi$. Calculating the coefficients a_n and b_n gives

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x \cos nx \, dx = 0,$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x \sin nx \, dx = -\frac{2}{n}\cos n\pi.$$

Hence,

$$x = 2[(-1/1 \cos \pi) \sin x + (-1/2 \cos 2\pi) \sin 2x + (-1/3 \cos 3\pi) \sin 3x + \cdots]$$

or

$$x = 2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots\right).$$

In this particular case, only the sine terms remain. It may be noted that whenever the function $f(x)$ is an odd function, that is, when $f(-x) = -f(x)$, then $a_n = 0$, for $n = 0, 1, 2 \cdots$, since, for such a function,

$$\int_{-\pi}^{0} f(x) \cos nx \, dx = -\int_{0}^{\pi} f(x) \cos nx \, dx.$$

Similarly, if $f(x)$ is an even function, that is, when $f(-x) = f(x)$ then $b_n = 0$, for $n = 1, 2, 3, \cdots$, since

$$\int_{-\pi}^{0} f(x) \sin nx \, dx = -\int_{0}^{\pi} f(x) \sin nx \, dx,$$

so that the function would be represented by a series of cosine terms.

If in the foregoing illustration the first four terms be plotted by composition of

$$y = 2 \sin x, \quad y = -\sin 2x, \quad y = 2/3 \sin 3x, \quad y = -1/2 \sin 4x,$$

the curve

$$y = 2 \sin x - \sin 2x + 2/3 \sin 3x - 1/2 \sin 4x$$

is obtained. It is represented on Fig. 10. As the number of terms is increased, the approximating curves approach $y = x$ as a limit for all values of x, $-\pi < x < \pi$, but not for $x = \pm\pi$. Since the series has period 2π, it represents the discontinuous function shown in Fig. 11 by a series of parallel lines. It should be noted that each term of the series is continuous and the function from which the series was derived is continuous, but the function represented by the series has finite discontinuities at

$x = \pm(2k + 1)\pi$. At such points the series converges to zero, which is one-half the value of the sum of the right- and left-hand limits.

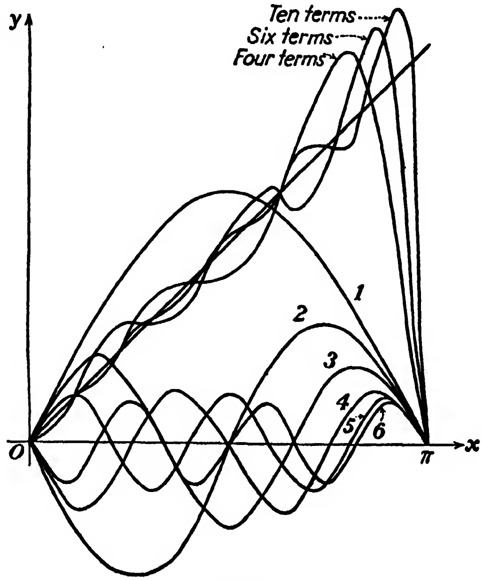

FIG. 10.

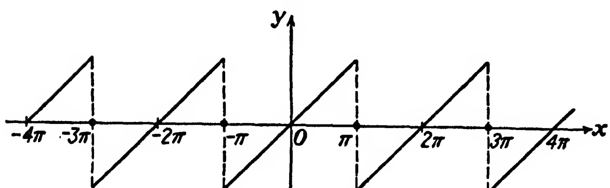

FIG. 11.

Illustrative Example 2. Develop $f(x)$ in Fourier series in the interval $(-\pi, \pi)$, if

$$f(x) = 0, \qquad \text{for } -\pi < x < 0,$$
$$= \pi, \qquad \text{for } \quad 0 < x < \pi.$$

Now

$$a_0 = \frac{1}{\pi}\left(\int_{-\pi}^{0} 0 \cdot dx + \int_0^{\pi} \pi\, dx\right) = \pi,$$

$$a_n = \frac{1}{\pi}\int_0^{\pi} \pi \cos nx\, dx = 0,$$

$$b_n = \frac{1}{\pi}\int_0^\pi \pi \sin nx\,dx = \frac{1}{n}(1 - \cos n\pi).$$

The series is then

$$\frac{\pi}{2} + 2\left(\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots\right).$$

The graph of $f(x)$ from $-\pi$ to π consists of the x-axis from $-\pi$ to 0, and the line AB from 0 to π (see Fig. 12). There is a finite discontinuity for $x = 0$. For $x = 0$ the series reduces

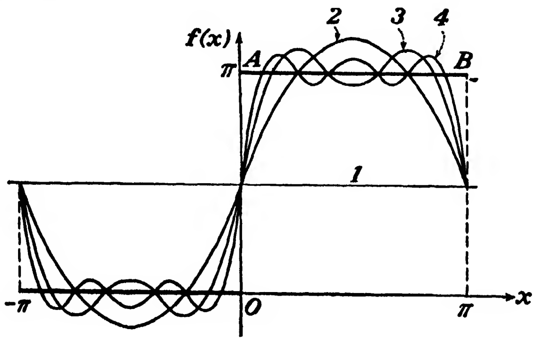

FIG. 12.

to $\pi/2$, which is equal to half the sum of $\lim\limits_{\epsilon\to 0} f(0 - \epsilon)$ and $\lim\limits_{\epsilon\to 0} f(0 + \epsilon)$. It may be observed from the series that every approximation curve will pass through the point $(0, \pi/2)$. The figure shows the first, second, third, and fourth approximation curves, whose equations are

$$y = \frac{\pi}{2}, \quad y = \frac{\pi}{2} + 2\sin x, \quad y = \frac{\pi}{2} + 2\left(\sin x + \frac{\sin 3x}{3}\right),$$

$$y = \frac{\pi}{2} + 2\left(\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5}\right),$$

as well as the graph of $f(x)$.

At $x = \pm\pi$ the series reduces to $\pi/2$, and again every approximation curve gives this same value for the ordinate at $\pm\pi$. This value is one-half the sum of $f(-\pi+)$ and $f(\pi-)$.

Illustrative Example 3. Let $f(x)$ be defined by the relations

$$\begin{aligned} f(x) &= -\pi, &&\text{if } -\pi < x < 0, \\ &= x, &&\text{if } \quad 0 < x < \pi; \end{aligned}$$

then the Fourier coefficients for $f(x)$ are given by

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dx = \frac{1}{\pi}\left(\int_{-\pi}^{0} -\pi\,dx + \int_{0}^{\pi} x\,dx\right)$$
$$= \frac{1}{\pi}\left(-\pi^2 + \frac{\pi^2}{2}\right) = -\frac{\pi}{2},$$
$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx = \frac{1}{\pi}\left(\int_{-\pi}^{0} -\pi\cos nx\,dx + \int_{0}^{\pi} x\cos nx\,dx\right)$$
$$= \frac{1}{\pi}\left(0 + \frac{\cos n\pi}{n^2} - \frac{1}{n^2}\right)$$
$$= \frac{1}{\pi}\left(\frac{\cos n\pi - 1}{n^2}\right),$$
$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx = \frac{1}{\pi}\left(\int_{-\pi}^{0} -\pi\sin nx\,dx + \int_{0}^{\pi} x\sin nx\,dx\right)$$
$$= \frac{1}{\pi}\left(\frac{\pi}{n} - \frac{\pi}{n}\cos n\pi - \frac{\pi}{n}\cos n\pi\right)$$
$$= \frac{1}{n}(1 - 2\cos n\pi).$$

Therefore,

$$f(x) = -\frac{\pi}{4} - \frac{2}{\pi}\cos x - \frac{2}{\pi}\frac{\cos 3x}{3^2} - \frac{2}{\pi}\frac{\cos 5x}{5^2} - \cdots$$
$$+ 3\sin x - \frac{\sin 2x}{2} + \frac{3\sin 3x}{3} - \frac{\sin 4x}{4} + \frac{3\sin 5x}{5} - \cdots.$$

When $x = 0$, the series reduces to

$$-\frac{\pi}{4} - \frac{2}{\pi}\left(\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right),$$

which must coincide with (see Fig. 13)

$$\frac{f(0+) + f(0-)}{2} = -\frac{\pi}{2}.$$

Thus,

$$-\frac{\pi}{4} - \frac{2}{\pi}\left(\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right) = -\frac{\pi}{2}.$$

Hence,

$$\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}.$$

Also, for $x = \pm\pi$, the series gives

$$-\frac{\pi}{4} + \frac{2}{\pi}\left(\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right) = 0,$$

since

$$\frac{f(-\pi+) + f(\pi-)}{2} = 0.$$

This example suggests the use of Fourier series in evaluating sums of series of constants.

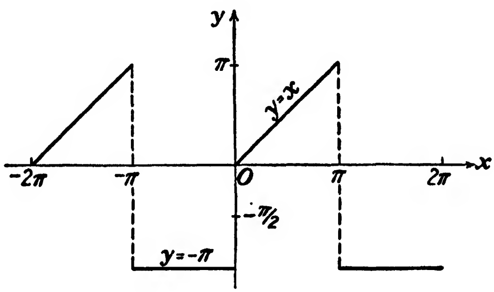

FIG. 13.

PROBLEMS

1. Show that

$$x^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty} (-1)^n \frac{\cos nx}{n^2}, \quad (-\pi \leq x \leq \pi).$$

2. If

$$\begin{aligned} f(x) &= -x && \text{for } -\pi < x < 0, \\ &= 0 && \text{for } \ 0 < x < \pi, \end{aligned}$$

then

$$f(x) = \frac{\pi}{4} - \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{\cos (2n-1)x}{(2n-1)^2} + \sum_{n=1}^{\infty} \frac{(-1)^n \sin nx}{n}.$$

3. If

$$\begin{aligned} f(x) &= 0 && \text{for } -\pi \leq x \leq 0, \\ &= \sin x && \text{for } \ 0 \leq x \leq \pi, \end{aligned}$$

then

$$f(x) = \frac{1}{\pi} - \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{\cos 2nx}{4n^2 - 1} + \frac{1}{2}\sin x.$$

4. If $f(x) = e^x$ in the interval $(0, 2\pi)$, then

$$e^x = \frac{e^{2\pi} - 1}{\pi}\left(\frac{1}{2} + \sum_{n=1}^{\infty} \frac{\cos nx}{1+n^2} - \sum_{n=1}^{\infty} \frac{n \sin nx}{1+n^2}\right).$$

5. Deduce from Prob. 1 that

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n^2} = \frac{\pi^2}{12}.$$

6. Show that

$$\cos \alpha x = \frac{\sin \pi\alpha}{\pi\alpha} + \sum_{n=1}^{\infty} (-1)^n \frac{2\alpha \sin \pi\alpha}{\pi(\alpha^2 - n^2)} \cos nx,$$

if $-\pi \leq x \leq \pi$.

7. Deduce from Prob. 6 that

$$\cot \pi\alpha = \frac{1}{\pi}\left(\frac{1}{\alpha} - \sum_{n=1}^{\infty} \frac{2\alpha}{n^2 - \alpha^2}\right).$$

8. Deduce from the expansion of $f(x) = x + x^2$ in Fourier series in the interval $(-\pi, \pi)$ that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

9. Expand $x \sin x$ and $x \cos x$ in Fourier series in the interval $(0, 2\pi)$.

10. Find the Fourier series expansion for $f(x)$, if

$$\begin{aligned} f(x) &= \frac{\pi}{2} \qquad \text{for } -\pi < x < \frac{\pi}{2}, \\ &= 0 \qquad \text{for } \frac{\pi}{2} < x < \pi. \end{aligned}$$

20. Sine and Cosine Series. The Fourier expansion for $f(x)$ in $(-\pi, \pi)$ has the form (19-1), in which the coefficients a_n and b_n are given by (19-2) and (19-3). As previously observed (Sec. 19), if $f(x)$ is an even function, (19-1) reduces to a series containing only cosine terms; and if $f(x)$ is an odd function, (19-1) reduces to a series containing only sine terms. Now suppose that it is desired that $f(x)$ be expanded in a Fourier series which will be used for the interval 0 to π only. In that case, it is frequently convenient to obtain the expansion in terms of sines alone or in terms of cosines alone. For this purpose, define

$$F(x) \equiv f(x) \qquad \text{for } 0 < x < \pi$$

and

$$F(x) \equiv f(-x) \qquad \text{for } -\pi < x < 0,$$

so that $F(x)$ is an even function identical with $f(x)$ in $0 < x < \pi$.

For an even function:

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(x) \sin nx\, dx \\ &= \frac{1}{\pi}\left[\int_{-\pi}^{0} F(x) \sin nx\, dx + \int_{0}^{\pi} F(x) \sin nx\, dx\right] \\ &= \frac{1}{\pi}\left[-\int_{0}^{\pi} F(-x) \sin (-nx)(-dx) + \int_{0}^{\pi} F(x) \sin nx\, dx\right] \\ &= \frac{1}{\pi}\left[-\int_{0}^{\pi} F(x) \sin nx\, dx + \int_{0}^{\pi} F(x) \sin nx\, dx\right] = 0, \end{aligned}$$

and

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(x) \cos nx\, dx \\ &= \frac{1}{\pi}\left[\int_{-\pi}^{0} F(x) \cos nx\, dx + \int_{0}^{\pi} F(x) \cos nx\, dx\right] \\ &= \frac{1}{\pi}\left[\int_{0}^{\pi} F(x) \cos nx\, dx + \int_{0}^{\pi} F(x) \cos nx\, dx\right] \\ &= \frac{2}{\pi}\int_{0}^{\pi} F(x) \cos nx\, dx. \end{aligned}$$

Hence, in the expansion of $F(x)$ in the interval $-\pi < x < \pi$, only the cosine terms appear. Moreover, $F(x)$ is identical with $f(x)$ for $0 < x < \pi$. Therefore,*

$$\text{(20-1)} \quad f(x) = \frac{a_0}{2} + a_1 \cos x + a_2 \cos 2x + \cdots + a_n \cos nx + \cdots$$

in the interval $(0, \pi)$, where

$$\text{(20-2)} \qquad a_n = \frac{2}{\pi}\int_{0}^{\pi} f(x) \cos nx\, dx.$$

Similarly, if $F(x)$ be defined so that

$$F(x) \equiv f(x) \qquad \text{for } 0 < x < \pi$$

and

$$F(x) \equiv -f(-x) \qquad \text{for } -\pi < x < 0,$$

* If $f(x)$ has a finite discontinuity at the point $x = x_0$, then the left-hand member of (20-1) is defined to be $\frac{1}{2}[f(x_0+) + f(x_0-)]$.

then the a_n all vanish and

(20-3) $f(x) = b_1 \sin x + b_2 \sin 2x + \cdots + b_n \sin nx + \cdots,$
where

$$b_n = \frac{2}{\pi} \int_0^\pi f(x) \sin nx \, dx. \tag{20-4}$$

Thus, $f(x)$ can be represented in the interval $0 < x < \pi$ by either (20-1) or (20-3). Frequently, one series is more desirable than the other.

Example. As has been determined already (see Illustrative Example 1, Sec. 19), the expansion for $f(x) = x$ in a sine series is

$$x = 2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots\right).$$

This series represents $f(x) = x$ in the interval $(-\pi, \pi)$. If one is inter-

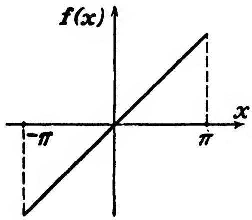

FIG. 14.

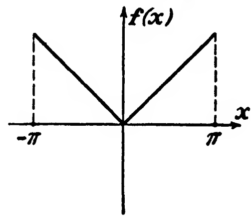

FIG. 15.

ested in the values of the function in the interval $(0, \pi)$, the same function can be expanded in a series of cosines.

In fact, in the interval $(0, \pi)$,

$$x = \frac{\pi}{2} - \frac{4}{\pi}\left(\frac{\cos x}{1} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \cdots\right),$$

since

$$a_0 = \frac{2}{\pi} \int_0^\pi x \, dx = \pi$$

and

$$a_n = \frac{2}{\pi} \int_0^\pi x \cos nx \, dx = \frac{2}{n^2 \pi} [(-1)^n - 1].$$

The sine series represents the odd function shown in Fig. 14 and the cosine series the even function in Fig. 15. The two graphs are identical in the interval $(0, \pi)$.

PROBLEMS

1. Show that if c is a constant, then, in $0 < x < \pi$,

$$c = c\frac{4}{\pi}\left(\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots\right).$$

2. Give sine and cosine developments of $y = x \sin x$ in the interval $(0, \pi)$.

3. Show that, in $(0, \pi)$,

$$x^2 = \frac{2}{\pi}\left[\left(\frac{\pi^2}{1} - \frac{4}{1^3}\right)\sin x - \frac{\pi^2}{2}\sin 2x + \left(\frac{\pi^2}{3} - \frac{4}{3^3}\right)\sin 3x - \frac{\pi^2}{4}\sin 4x + \cdots\right].$$

4. Prove that, if $f(x)$ is any function of x, it can be expressed as the sum of an even function of x and an odd function of x.

5. Show that, if $f(x) = x$ for $0 < x < \pi/2$ and $f(x) = \pi - x$ for $\pi/2 < x < \pi$, then

$$f(x) = \frac{\pi}{4} - \frac{2}{\pi}\left(\frac{\cos 2x}{1^2} + \frac{\cos 6x}{3^2} + \frac{\cos 10x}{5^2} + \cdots\right).$$

6. Show that

$$\log\left(2\sin\frac{x}{2}\right) = -\sum_{n=1}^{\infty}\frac{\cos nx}{n}, \qquad \text{if } 0 < x < \pi.$$

7. Find the expansion in the series of sines, if

$$\begin{aligned} f(x) &= \frac{\pi}{4}x, && \text{for } 0 \leq x \leq \frac{\pi}{2}, \\ &= \frac{\pi}{4}(\pi - x), && \text{for } \frac{\pi}{2} \leq x \leq \pi. \end{aligned}$$

8. Expand $f(x) = e^x$ in the series of cosines in the interval $(0, \pi)$.

21. Extension of Interval of Expansion. The methods developed up to this point restrict the interval in which $f(x)$ can be expanded in a Fourier series to $(-\pi, \pi)$. In many problems, it is desirable to develop $f(x)$ in a Fourier series that will be valid over a wider interval. In order to obtain an expansion that will hold for the interval $(-l, l)$, change the variable by replacing x by $\frac{l}{\pi}z$. Then $f(x) = f\left(\frac{l}{\pi}z\right)$ can be developed in a Fourier series in z,

$$f\left(\frac{l}{\pi}z\right) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nz + \sum_{n=1}^{\infty} b_n \sin nz, \tag{21-1}$$

in which

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{lz}{\pi}\right)\cos nz\, dz$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{lz}{\pi}\right) \sin nz\, dz.$$

The expression (21-1) will be valid for $-\pi < z < \pi$; but $z = \pi x/l$ so that (21-1) becomes

$$(21\text{-}2) \qquad f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{l} + \sum_{n=1}^{\infty} b_n \sin\frac{n\pi x}{l}.$$

Also,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{lz}{\pi}\right) \cos nz\, dz = \frac{1}{l}\int_{-l}^{l} f(x) \cos\frac{n\pi x}{l}\, dx$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{lz}{\pi}\right) \sin nz\, dz = \frac{1}{l}\int_{-l}^{l} f(x) \sin\frac{n\pi x}{l}\, dx.$$

Example. Develop $f(x)$ in Fourier series in the interval $(-2, 2)$, if $f(x) = 0$ for $-2 < x < 0$ and $f(x) = 1$ for $0 < x < 2$. Here

$$a_0 = \frac{1}{2}\left(\int_{-2}^{0} 0 \cdot dx + \int_0^2 1 \cdot dx\right) = 1,$$

$$a_n = \frac{1}{2}\left(\int_{-2}^{0} 0 \cdot \cos\frac{n\pi x}{2}\, dx + \int_0^2 1 \cdot \cos\frac{n\pi x}{2}\, dx\right) = \frac{1}{n\pi} \sin\frac{n\pi x}{2}\bigg|_0^2 = 0,$$

$$b_n = \frac{1}{2}\left(\int_{-2}^{0} 0 \cdot \sin\frac{n\pi x}{2}\, dx + \int_0^2 1 \cdot \sin\frac{n\pi x}{2}\, dx\right) = \frac{1}{n\pi}(1 - \cos n\pi).$$

Therefore,

$$f(x) = \frac{1}{2} + \frac{2}{\pi}\left(\sin\frac{\pi x}{2} + \frac{1}{3}\sin\frac{3\pi x}{2} + \frac{1}{5}\sin\frac{5\pi x}{2} + \cdots\right).$$

PROBLEMS

1. The expansion of $f(x)$ is desired for $0 < x < l$. If $F(x) \equiv f(x)$ for $0 < x < l$ and $F(x) \equiv -f(-x)$ for $-l < x < 0$ [that is, $F(x)$ is defined as an odd function], show that the expansion of $F(x)$ and $f(x)$ for $0 < x < l$ is

$$\sum_{n=1}^{\infty} b_n \sin\frac{n\pi x}{l},$$

where

$$b_n = \frac{2}{l}\int_0^l f(x) \sin\frac{n\pi x}{l}\, dx.$$

If $\varphi(x) \equiv f(x)$ for $0 < x < l$ and $\varphi(x) \equiv f(-x)$ for $-l < x < 0$ [that is, $\varphi(x)$ is defined as an even function], show that the expansion of $\varphi(x)$

and $f(x)$ for $0 < x < l$ is

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{l},$$

where

$$a_n = \frac{2}{l}\int_0^l f(x) \cos \frac{n\pi x}{l}\, dx.$$

2. Using the results of the preceding problem, obtain the sine and cosine expansions of the following functions:

(a) $f(x) = 1$ in the interval $(0, 2)$;
(b) $f(x) = x$ in the interval $(0, 1)$;
(c) $f(x) = x^2$ in the interval $(0, 3)$.

3. Expand $f(x) = \cos \pi x$ in the interval $(-1, 1)$.

4. Expand

$$\begin{aligned} f(x) &= \tfrac{1}{4} - x, && \text{if } 0 < x < \tfrac{1}{2}, \\ &= x - \tfrac{3}{4}, && \text{if } \tfrac{1}{2} < x < 1, \end{aligned}$$

in the series of sines.

5. Find the expansion in the series of cosines, if

$$\begin{aligned} f(x) &= 0, && \text{if } 0 < x < 1, \\ &= 1, && \text{if } 1 < x < 2. \end{aligned}$$

6. Expand $f(x) = |x|$ in the series of cosines in the interval $(-1, 1)$.

7. Show that the series

$$\frac{l}{\pi}\sum_{n=1}^{\infty} \frac{1}{n} \sin \frac{2n\pi x}{l}$$

represents $\tfrac{1}{2}l - x$ when $0 < x < l$.

8. Find the expansion in the series of cosines, if

$$\begin{aligned} f(x) &= 1 && \text{when } 0 < x < \pi, \\ &= 0 && \text{when } \pi < x < 2\pi. \end{aligned}$$

22. Complex Form of Fourier Series. The Fourier series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{22-1}$$

with

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \cos nt\, dt, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \sin nt\, dt,$$

can be written, with the aid of the Euler formula*

$$e^{iu} = \cos u + i \sin u, \tag{22-2}$$

* See Sec. 73.

in an equivalent form, namely,

$$f(x) = \sum_{n=-\infty}^{n=+\infty} c_n e^{inx}, \tag{22-3}$$

where the coefficients c_n are defined by the equation

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt. \tag{22-4}$$

The index of summation n in (22-3) runs through the set of all positive and negative integral values including zero.

The equivalence of (22-3) and (22-1) can be established in the following manner: Substituting from (22-2) in (22-4) gives, for $n > 0$,

$$\begin{aligned} c_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,(\cos nt - i\sin nt)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt - \frac{i}{2\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt \\ &= \frac{a_n}{2} - i\frac{b_n}{2}. \end{aligned}$$

A similar calculation for $n < 0$ gives

$$c_{-n} = \frac{a_n}{2} + i\frac{b_n}{2},$$

while

$$c_0 = \frac{a_0}{2}.$$

Now (22-3) can be written in the form

$$f(x) = c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + \sum_{n=1}^{\infty} c_{-n}e^{-inx}.$$

Making use of the expressions for the c_n just found gives

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty}\frac{a_n - ib_n}{2}e^{inx} + \sum_{n=1}^{\infty}\frac{a_n + ib_n}{2}e^{-inx} \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\frac{e^{inx} + e^{-inx}}{2} - i\sum_{n=1}^{\infty} b_n\frac{e^{inx} - e^{-inx}}{2}. \end{aligned}$$

Recalling that

$$e^{iu} + e^{-iu} = 2 \cos u \qquad \text{and} \qquad e^{iu} - e^{-iu} = 2i \sin u$$

gives

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

which establishes the identity of (22-3) with (22-1).

PROBLEM

Show that the Fourier series in the interval $(-l, l)$ can be written in the form

$$f(x) = \sum_{n=-\infty}^{n=+\infty} c_n e^{\frac{in\pi x}{l}}$$

where

$$c_n = \frac{1}{2l} \int_{-l}^{l} f(t) e^{-\frac{in\pi t}{l}}\, dt.$$

23. Differentiation and Integration of Fourier Series. Some general rules concerning differentiation and integration of infinite series were given in Sec. 8. It may be added here that, if the Fourier series represents a function $f(x)$, then the term-by-term integral of the series will converge to the integral of $f(x)$. Thus, in Sec. 20 it was found that the function x can be expanded in a cosine series in $(0, \pi)$ to give

$$\text{(23-1)} \quad x = \frac{\pi}{2} - \frac{4}{\pi}\left(\frac{\cos x}{1^2} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \cdots\right).$$

The term-by-term integral of this series gives

$$\text{(23-2)} \quad \frac{x^2}{2} = \int_0^x \frac{\pi}{2}\, dx - \frac{4}{\pi}\left(\int_0^x \frac{\cos x}{1^2}\, dx + \int_0^x \frac{\cos 3x}{3^2}\, dx + \int_0^x \frac{\cos 5x}{5^2}\, dx + \cdots\right) = \frac{\pi}{2} x - \frac{4}{\pi}\left(\frac{\sin x}{1^3} + \frac{\sin 3x}{3^3} + \frac{\sin 5x}{5^3} + \cdots\right).$$

On account of the presence of the term $\pi x/2$ in the right member of this equation, the resulting series is not a Fourier series. However, if the sine series development for x (Illustrative Example 1, Sec. 19) be substituted in this term and the like terms collected,

the resulting series will be the development of $x^2/2$ in $(0, \pi)$ in a sine series.

The series resulting from term-by-term differentiation of a Fourier series converges more slowly than the original series, and, in fact, it may diverge. Thus, term-by-term differentiation of

$$x = 2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots\right)$$

gives the series

$$2(\cos x - \cos 2x + \cos 3x - \cdots),$$

which is divergent, as was observed in Sec. 8. Accordingly, great caution must be exercised in differentiating Fourier series termwise.*

24. Orthogonal Functions. A set of continuous functions

$$u_1(x),\ u_2(x),\ \cdots,\ u_n(x),\ \cdots, \tag{24-1}$$

which do not vanish identically in the interval $a \leq x \leq b$, is said to be *orthogonal* with respect to the interval (a, b) if the functions $u_i(x)$ satisfy the relations

$$\int_a^b u_i(x)u_j(x)\,dx = 0, \qquad \text{if } i \neq j. \tag{24-2}$$

For $i = j$, (24-2) becomes

$$\int_a^b [u_i(x)]^2\,dx \equiv c_i^2,$$

where c_i^2 certainly is not zero.

If each of the orthogonal functions $u_i(x)$ be divided by c_i, there will be obtained a system of *normal* orthogonal functions,

$$v_1(x) = \frac{u_1(x)}{c_1}, \qquad v_2(x) = \frac{u_2(x)}{c_2}, \cdots, \qquad v_n(x) = \frac{u_n(x)}{c_n}, \cdots,$$

that are characterized by the property that

$$\begin{aligned} \int_a^b v_i(x)v_j(x)\,dx &= 0, \qquad \text{if } i \neq j, \\ &= 1, \qquad \text{if } i = j. \end{aligned} \tag{24-3}$$

Consider a set of normal orthogonal functions $v_i(x)$, and assume that an arbitrary function $f(x)$ can be expanded in a series

* Some important theorems in regard to this will be found in I. S. Sokolnikoff, Advanced Calculus, Sec. 108.

$$(24\text{-}4) \quad f(x) = a_1v_1(x) + a_2v_2(x) + \cdots + a_nv_n(x) + \cdots$$
$$\equiv \sum_{i=1}^{\infty} a_iv_i(x),$$

which can be integrated term by term. Multiplying both sides of (24-4) by $v_j(x)$ and integrating term by term between the limits a and b yield

$$\int_a^b f(x)v_j(x)\,dx = \sum_{i=1}^{\infty} a_i \int_a^b v_i(x)v_j(x)\,dx,$$

which, by virtue of (24-3), gives the formula

$$(24\text{-}5) \qquad a_i = \int_a^b f(x)v_i(x)\,dx, \qquad (i = 1, 2, 3, \cdots).$$

The numbers a_i are known as the Fourier coefficients of the function $f(x)$ associated with the system of normal and orthogonal functions

$$v_1(x), v_2(x), \cdots, v_n(x), \cdots.$$

The set of functions

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \cdots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \cdots$$

is obviously a normal orthogonal set in the interval $(-\pi, \pi)$, and the development of a function $f(x)$ with the aid of this particular set of orthogonal functions is precisely the Fourier development of $f(x)$. Among many other useful sets of orthogonal functions are the functions of Bessel and Legendre, which are of frequent occurrence in applied mathematics and will be used in the discussion of some important problems in Secs. 101, 113, and 114.

CHAPTER III

SOLUTION OF EQUATIONS

Students of engineering, physics, chemistry, and other sciences meet the problem of the solution of equations at every stage of their work. This chapter gives a brief outline of some of the algebraic, graphical, and numerical methods of obtaining the real roots of equations with real coefficients, of types that occur frequently in the applied sciences. It also contains a short summary of those parts of the theory of determinants and the theory of matrices that are immediately applicable to the solution of systems of linear equations.

25. Graphical Solutions. The subject of the solution of equations will be introduced by considering a simple problem that any engineer may be called upon to solve.

It is required to design a hollow cast-iron sphere, 1 in. in thickness, that will just float in water. It is assumed that the air in the cavity is completely exhausted. The specific gravity of cast iron will be denoted by ρ, for convenience.

By the law of Archimedes, the weight of the sphere must equal the weight of the displaced water. This gives the condition on the radius of the sphere, namely,

$$\tfrac{4}{3}\pi x^3 = \tfrac{4}{3}\pi\rho[x^3 - (x - 1)^3].$$

Simplifying gives

$$x^3 - 3\rho x^2 + 3\rho x - \rho = 0. \tag{25-1}$$

It will be convenient to remove the second-degree term in (25-1). To accomplish this, let $x = y + k$, giving

$$y^3 + 3y^2k + 3yk^2 + k^3 - 3\rho(y^2 + 2yk + k^2) + 3\rho(y + k) - \rho = 0,$$

or

$$y^3 + (3k - 3\rho)y^2 + (3k^2 - 6\rho k + 3\rho)y + k^3 - 3\rho k^2 + 3\rho k - \rho = 0.$$

Choosing $k = \rho$ makes the equation reduce to

$$y^3 + (3\rho - 3\rho^2)y - 2\rho^3 + 3\rho^2 - \rho = 0. \tag{25-2}$$

For cast iron, $\rho = 7.5$, and (25-2) becomes

$$y^3 - 146.25y - 682.5 = 0. \tag{25-3}$$

If (25-3) is solved, the solution of (25-1) is also determined, since $x = y + 7.5$.

A graphical method of solution will be used. The solution of (25-3) is equivalent to the simultaneous solution of the system

$$\begin{cases} z = y^3, \\ z = 146.25y + 682.5. \end{cases} \tag{25-4}$$

The accompanying figure (Fig. 16) represents the graphs of the two functions of (25-4); since they intersect at $y = 14.0$, this value gives an approximate solution of (25-3). The corresponding solution of (25-1) is $x = 21.5$. From the graph, it is clear that there is only one real solution of (25-4) and hence of (25-3).

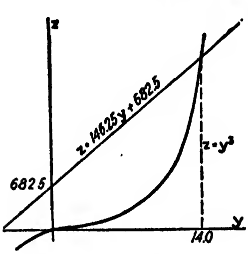

Fig. 16.

This graphical method can be applied to any cubic equation. The general fourth-degree equation (quartic) can also be reduced to a form that is convenient for graphical methods of solution.

Consider the quartic

$$x^4 + ax^3 + bx^2 + cx + d = 0.$$

Let $x = y + k$, as in the cubic equation. This substitution gives

$$y^4 + y^3(4k + a) + y^2(6k^2 + 3ak + b) \\ + y(4k^3 + 3ak^2 + 2bk + c) + k^4 + ak^3 + bk^2 + ck + d = 0.$$

In order to remove the term in y^3, choose $k = -\dfrac{a}{4}\cdot$ This reduces the equation to the form

$$y^4 + Ay^2 + By + C = 0.$$

If $A > 0$, the further transformation $y = \sqrt{A}\, z$ is made, and the equation is reduced to

$$A^2z^4 + A^2z^2 + B\sqrt{A}\, z + C = 0,$$

or

$$z^4 + z^2 + pz + q = 0.$$

The solutions of this equation are the same as the solutions of the simultaneous system

$$u = z^4 + z^2,$$
$$u = -pz - q.$$

The graphs of these two functions are easily plotted, and the solutions can be read from the graph. In case $A < 0$, the transformation would be $y = \sqrt{-A}\,z$, which leads to the equation

$$z^4 - z^2 + pz + q = 0$$

and the graphical solution of the system

$$u = z^4 - z^2,$$
$$u = -pz - q.$$

This method of solution for the real roots of an equation is also applicable to many transcendental equations. In order to solve

$$Ax - B \sin x = 0,$$

write it as

$$ax - \sin x = 0,$$

and plot the curves of the simultaneous system

$$y = \sin x,$$
$$y = ax.$$

Similarly, the equation

$$a^x - x^2 = 0$$

can be solved graphically by plotting the curves of the equivalent simultaneous system

$$y = a^x,$$
$$y = x^2.$$

PROBLEMS

1. Solve graphically

(a) $2^x - x^2 = 0,$
(b) $x^4 - x - 1 = 0,$
(c) $x^5 - x - 0.5 = 0,$
(d) $e^x + x = 0.$

2. Find, graphically, the root of

$$\tan x - x = 0$$

nearest $\frac{3}{2}\pi$.

26. Algebraic Solution of Cubic. The graphical method of solution is perfectly general, but its accuracy depends upon the accurate construction of the graphs of the equations in the simultaneous systems. This is often extremely laborious and, at most, yields only an approximate value of the roots.

In the case of the linear equation $ax + b = 0$, where $a \neq 0$, the solution is $x = -b/a$. For the quadratic equation $ax^2 + bx + c = 0$, where $a \neq 0$, there are two solutions given by $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$.

The question naturally arises as to the possibility of obtaining expressions for the roots of algebraic equations of degree higher than 2. This section will be devoted to a derivation of the solutions of the general cubic equation

$$a_0x^3 + a_1x^2 + a_2x + a_3 = 0, \qquad a_0 \neq 0.$$

Dividing through by a_0 gives

$$x^3 + bx^2 + cx + d = 0, \tag{26-1}$$

and the x^2 term can be removed by making the change of variable

$$x = y - \frac{b}{3}.$$

The resulting equation is

$$y^3 + py + q = 0, \tag{26-2}$$

where

$$p = c - \frac{b^2}{3}$$

and

$$q = d - \frac{bc}{3} + \frac{2b^3}{27}.$$

In order to solve (26-2), assume that

$$y = A + B, \tag{26-3}$$

so that

$$y^3 = A^3 + B^3 + 3AB(A + B).$$

Substitute in this last equation for $A + B$, from (26-3), and there is obtained the equation

$$y^3 - 3ABy - (A^3 + B^3) = 0. \tag{26-4}$$

A comparison of (26-4) with (26-2) shows that

$$3AB = -p \qquad \text{and} \qquad A^3 + B^3 = -q,$$

or

$$A^3B^3 = -\frac{p^3}{27} \quad \text{and} \quad A^3 + B^3 = -q. \tag{26-5}$$

If B^3 is eliminated by substituting from the second of Eqs. (26-5) into the first, there appears the quadratic equation in A^3,

$$(A^3)^2 + qA^3 - \frac{p^3}{27} = 0,$$

whose roots are

$$A^3 = \frac{-q \pm \sqrt{q^2 + \frac{4p^3}{27}}}{2}.$$

The solution for B^3 yields precisely the same values. However, in order to satisfy Eq. (26-5), choose*

$$\left\{\begin{aligned} A^3 &= \frac{-q + \sqrt{q^2 + \frac{4p^3}{27}}}{2}, \\ B^3 &= \frac{-q - \sqrt{q^2 + \frac{4p^3}{27}}}{2}. \end{aligned}\right. \tag{26-6}$$

If the values of y are to be determined from (26-3), it is necessary to find the cube roots of A^3 and B^3. Recall that if $x^3 = a^3$, then the solutions for x are given by a, ωa, and $\omega^2 a$, where $\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\omega^2 = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$ are the complex roots of unity. Hence, if one cube root of A^3 be denoted by α and one cube root of B^3 by β, the cube roots of A^3 are

$$\alpha, \qquad \omega\alpha, \qquad \text{and} \qquad \omega^2\alpha,$$

whereas those of B^3 are

$$\beta, \qquad \omega\beta, \qquad \text{and} \qquad \omega^2\beta.$$

It would appear that there are nine choices for y, but it should be remembered that the values must be paired so that $3AB = -p$. The only pairs that satisfy this condition are α and β, $\omega\alpha$ and $\omega^2\beta$, and $\omega^2\alpha$ and $\omega\beta$. Hence, the values of y are

$$y_1 = \alpha + \beta, \qquad y_2 = \omega\alpha + \omega^2\beta, \qquad y_3 = \omega^2\alpha + \omega\beta, \tag{26-7}$$

where

$$\alpha = \sqrt[3]{\frac{-q + \sqrt{q^2 + \frac{4p^3}{27}}}{2}} \quad \text{and} \quad \beta = \sqrt[3]{\frac{-q - \sqrt{q^2 + \frac{4p^3}{27}}}{2}}.$$

* The opposite choice for the values of A^3 and B^3 simply interchanges their role in what follows.

The solutions of (26-1) can be obtained from the values given in (26-7) by recalling that $x = y - b/3$.

The expressions for α and β are quite complicated, and when the quantity under the square-root sign has a negative value the values of α and β cannot, in general, be determined. This is the so-called irreducible case of the cubic, which can, however, be solved by using a trigonometric method. This method will be described later in the section, but first it is important to find a criterion that will determine in advance which method should be used.

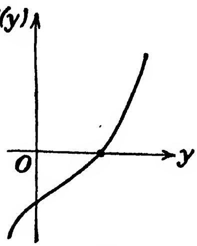

FIG. 17.

In order to determine the character of the roots of (26-2), whose coefficients are assumed to be real, consider the function

$$f(y) \equiv y^3 + py + q$$

and its maximum and minimum values. Since

$$f'(y) = 3y^2 + p,$$

it appears that, if $p > 0$, then $f'(y)$ is always positive and $f(y)$ is an increasing function. In this case the graph of $f(y)$ has the form shown in Fig. 17, and there is evidently only one real value for which $f(y) = 0$.

If $p < 0$, however, $f'(y)$ is zero when $y = \pm\sqrt{-p/3}$. Since $f''(y) = 6y$, it follows that $y = +\sqrt{-p/3}$ gives a minimum value to $f(y)$, whereas $y = -\sqrt{-p/3}$ furnishes a maximum value. The corresponding values of $f(y)$ are

$$q + \frac{2}{3}p\sqrt{-\frac{p}{3}}$$

and

$$q - \frac{2}{3}p\sqrt{-\frac{p}{3}}.$$

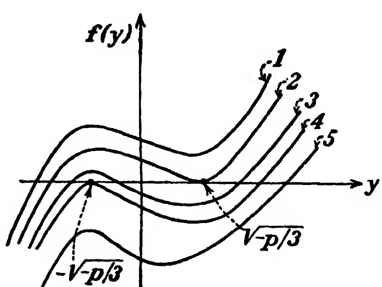

FIG. 18.

The graph of $f(y)$ will have the appearance of one of the curves in Fig. 18.

It is evident that $f(y) = 0$ will have only one real root if the graph of $f(y)$ has the appearance shown by (1) or (5), that is, if the maximum and minimum values of $f(y)$ are of the same sign. Hence,

$$\left(q + \frac{2}{3}p\sqrt{-\frac{p}{3}}\right)\left(q - \frac{2}{3}p\sqrt{-\frac{p}{3}}\right) > 0,$$

or

$$q^2 + \tfrac{4}{27}p^3 > 0,$$

is the condition that (26-2) have only one real root. It may be observed that this condition is automatically satisfied if $p \geqq 0$. It should be noted that, if $p = 0$, Eq. (26-2) reduces to $y^3 + q = 0$, which obviously has only one real root.

If (26-2) has three real and distinct roots, then the graph of $f(y)$ must have the appearance shown in (3), and it follows that the maximum and minimum values must be of opposite sign. Hence,

$$q^2 + \tfrac{4}{27}\, p^3 < 0$$

is the condition for three real and unequal roots.

If $q^2 + \tfrac{4}{27}p^3 = 0$, either the maximum or the minimum value of $f(y)$ must be zero [see (2) and (4)], and (26-2) will have three real roots, of which two will be equal (a so-called double root).

The expression

$$\Delta \equiv -27q^2 - 4p^3 \tag{26-8}$$

is called the discriminant of the cubic equation (26-2), for its value determines the character of the roots of the equation. The discriminant for (26-1), obtained by replacing p and q in (26-8) by their values in terms of b, c, and d, is

$$\Delta \equiv 18bcd - 4b^3d + b^2c^2 - 4c^3 - 27d^2. \tag{26-9}$$

It may be worth noting that the discriminant of any algebraic equation, with leading coefficient unity, is the product of the squares of the differences of the roots taken two at a time. Inasmuch as

$$(x_1 - x_2)^2(x_2 - x_3)^2(x_3 - x_1)^2 = (y_1 - y_2)^2(y_2 - y_3)^2(y_3 - y_1)^2,$$

the discriminant has the same value for (26-1) and (26-2).

In view of the definition of Δ, it follows that

if $\Delta < 0$, one root is real and two are complex;
if $\Delta = 0$, all the roots are real and two are equal;
if $\Delta > 0$, the three roots are real and unequal.

Example. Consider the cubic equation

$$x^3 + 3x^2 + 9x - 1 = 0.$$

From (26-9), it follows that $\Delta = -2592$, and hence there will be one real root and two complex roots. Setting $x = y - 1$ yields the reduced cubic

$$y^3 + 6y - 8 = 0,$$

and substituting $p = 6$ and $q = -8$ in (26-6) gives $A^3 = 4 + 2\sqrt{6}$ and $B^3 = 4 - 2\sqrt{6}$. Therefore, the solutions for y are

$$\sqrt[3]{4 + 2\sqrt{6}} + \sqrt[3]{4 - 2\sqrt{6}}, \qquad \omega\sqrt[3]{4 + 2\sqrt{6}} + \omega^2\sqrt[3]{4 - 2\sqrt{6}},$$

$$\text{and} \qquad \omega^2\sqrt[3]{4 + 2\sqrt{6}} + \omega\sqrt[3]{4 - 2\sqrt{6}}.$$

The solutions of the original equation can now be obtained by recalling that $x = y - 1$.

The discussion of the solution of the cubic equation will be concluded by giving the derivation of the expressions for the roots in the case where the roots are real and unequal (that is, when $\Delta \equiv -27q^2 - 4p^3 > 0$).

Let

$$-\frac{q}{2} = r\cos\theta$$

and

$$\sqrt{-\left(\frac{p^3}{27}+\frac{q^2}{4}\right)} = r\sin\theta.$$

Then*

$$\alpha = (r\cos\theta + ir\sin\theta)^{1/3} = r^{1/3}\left(\cos\frac{\theta}{3} + i\sin\frac{\theta}{3}\right)$$

and

$$\beta = (r\cos\theta - ir\sin\theta)^{1/3} = r^{1/3}\left(\cos\frac{\theta}{3} - i\sin\frac{\theta}{3}\right).$$

If it is noted that

$$\omega = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}$$

and

$$\omega^2 = \cos\frac{2\pi}{3} - i\sin\frac{2\pi}{3},$$

it is easily checked that the expressions for

$$y_1 \equiv \alpha + \beta, \qquad y_2 \equiv \omega\alpha + \omega^2\beta, \qquad y_3 \equiv \omega^2\alpha + \omega\beta$$

become

$$y_1 = 2r^{1/3}\cos\frac{\theta}{3}, \qquad y_2 = 2r^{1/3}\cos\frac{\theta+2\pi}{3}, \qquad y_3 = 2r^{1/3}\cos\frac{\theta+4\pi}{3}. \tag{26-10}$$

Since

$$r = \sqrt{-\frac{p^3}{27}}$$

and

$$\cos\theta = -\frac{q}{2}\sqrt{-\frac{27}{p^3}},$$

the values of y_1, y_2, and y_3 can be obtained directly from the coefficients of (26-2) or (26-1).

* By De Moivre's theorem $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$.

Example. Determine the real roots of

$$x^3 - 3x^2 + 3 = 0.$$

Here

$$\Delta = -4(-27)(3) - 27(9) > 0,$$

and the roots are all real and unequal. Since $p = -3$ and $q = 1$, it follows that $r = 1$ and $\cos \theta = -\frac{1}{2}$. Hence,

$$\theta = \frac{2\pi}{3},$$

and

$$y_1 = 2 \cos \frac{2\pi}{9}, \qquad y_2 = 2 \cos \frac{8\pi}{9}, \qquad y_3 = 2 \cos \frac{14\pi}{9}.$$

The solutions of the general quartic equation

$$x^4 + bx^3 + cx^2 + dx + e = 0$$

can be found, but the methods of obtaining the expressions for the roots depend upon the solution of an auxiliary cubic equation. Moreover, these expressions are, in general, so involved that they are practically useless for computation.* It has been shown that the ordinary operations of algebra are, in general, insufficient for the purpose of obtaining exact solutions of algebraic equations of degree higher than 4. However, it is possible to obtain the expressions for the solutions of the general equation of the fifth degree with the aid of elliptic integrals.

The reader should not confuse the problem of obtaining expressions for the exact solutions of the general algebraic equation with that of calculating numerical approximations to the roots of specific equations which have numerical coefficients. The latter problem will be discussed in Secs. 28 and 29, and it will be shown that the real roots of such equations can be computed to any desired degree of accuracy. Moreover, if the roots are rational they can always be determined exactly.

PROBLEMS

Determine the roots of the following equations:

(*a*) $y^3 - 2y - 1 = 0$;
(*b*) $y^3 - 146.25y - 682.5 = 0$;
(*c*) $x^3 - x^2 - 5x - 3 = 0$;
(*d*) $x^3 - 2x^2 - x + 2 = 0$;
(*e*) $x^3 - 6x^2 + 6x - 2 = 0$;
(*f*) $x^3 + 6x^2 + 3x + 18 = 0$;
(*g*) $2x^3 + 3x^2 + 3x + 1 = 0$.

* See Dickson, L. E., First Course in Theory of Equations, pp. 50–54; Burnside, W. S., and A. W. Panton, Theory of Equations, vol. 1, pp. 121–142.

27. Some Algebraic Theorems. The student of any applied science is usually interested in obtaining numerical values, correct to a certain number of decimal places, for the roots of equations. Unless the roots are rational, the expressions for the exact roots, provided that they can be found at all, are usually complicated and the process of determining numerical values from them is tedious. Accordingly, it is distinctly useful to consider other methods of finding these numerical values. Horner's method, Newton's method, and the method of interpolation are the ones most frequently used; they will be discussed in Secs. 28 and 29. However, all these methods are based on the assumption that a root has first been isolated, that is, that there have been determined two values of the variable such that between them lies one and only one root. In many practical problems the physical setup is a guide in this isolation process. This section contains a review of some theorems* from the theory of equations that provide preliminary information as to the character and location of the roots.

THEOREM 1. (Fundamental Theorem of Algebra.) *Every algebraic equation*

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0$$

has a root.

It should be noted that this theorem does not hold for non-algebraic equations. For example, the equation $e^x = 0$ has no root.

THEOREM 2. (Remainder Theorem.) *If the polynomial*

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$

*is divided by $x - b$ until the remainder is independent of x, then this remainder has the value $f(b)$.**

THEOREM 3. (Factor Theorem.) *If $f(b) = 0$, then $x - b$ is a factor of the polynomial $f(x)$ and b is a root of $f(x) = 0$.*

This theorem follows directly from Theorem 2. In many cases the easiest way to compute the value of $f(b)$ is to perform the division of $f(x)$ by $x - b$. This is a particularly useful

* Those students who are not already familiar with these theorems and their proofs will benefit by referring to H. B. Fine, College Algebra, pp. 425–453, and L. E. Dickson, First Course in the Theory of Equations, Chap. II.

method when the factor theorem is being used for the purpose of determining the roots of $f(x) = 0$. For if $x - b$ is a factor of $f(x)$, it follows that $f(x) = (x - b)\, g(x)$, where $g(x)$ is a polynomial of degree one less than that of $f(x)$. Obviously the roots of $g(x) = 0$ will be the remaining roots of $f(x) = 0$, so that only $g(x) = 0$ need be considered in attempting to find these roots. Moreover, when $f(x)$ is divided by $x - b$ the quotient is $g(x)$. If synthetic division is used, the computation is usually quite simple.

Example. If $f(x) = x^3 + 2x^2 + 2x + 1$ is divided by $x + 1$, the quotient is $x^2 + x + 1$ and the remainder is zero. Hence, $x = -1$ is a root of $f(x) = 0$ and the remaining roots are determined by solving $x^2 + x + 1 = 0$.

THEOREM 4. *Every algebraic equation of degree n has exactly n roots if a root of multiplicity m is counted as m roots.*

A root b of $f(x) = 0$ is said to be a root of multiplicity m if $(x - b)^m$ is a factor of $f(x)$ but $(x - b)^{m+1}$ is not a factor of $f(x)$.

It follows from Theorems 3 and 4 that the polynomial of degree n can be factored into n linear factors, so that

$$\begin{aligned} f(x) &\equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \\ &= a_0(x - x_1)(x - x_2) \cdots (x - x_n). \end{aligned}$$

THEOREM 5. *If*

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$

has integral coefficients and if $f(x) = 0$ has the rational root b/c, where b and c are integers without a common divisor, then b is an exact divisor of a_n and c is an exact divisor of a_0.

Example. Consider the equation

$$f(x) \equiv 2x^3 + x^2 + x - 1 = 0.$$

The only possible rational roots are ± 1 and $\pm \frac{1}{2}$. Since $f(1) = 3$, $f(-1) = -3$, $f(-\frac{1}{2}) = -\frac{3}{2}$, and $f(\frac{1}{2}) = 0$, it follows that $\frac{1}{2}$ is the only rational root. As a matter of fact, if $f(x)$ is divided by $x - \frac{1}{2}$ the quotient is $2x^2 + 2x + 2$ whose factors are 2, $x - \omega$, and $x - \omega^2$, where ω and ω^2 are the complex roots of unity.*

THEOREM 6. *Given $f(x) \equiv x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0$. If $f(a)$ and $f(b)$ are of opposite sign, then there exists at least*

* See Sec. 26 and the example following Theorem 9 of this section.

one root of $f(x) = 0$ between a and b. Moreover, the number of such roots is odd.

Graphically this means that $y = f(x)$ must cross the x-axis an odd number of times between a and b.

Example. If $f(x) \equiv 8x^3 - 12x^2 - 2x + 3 = 0$,

$$f(-1) = -15, \qquad f(0) = 3; \qquad f(1) = -3, \qquad f(2) = 15.$$

Since $f(-1)$ is negative and $f(0)$ is positive, there is at least one root between -1 and 0. Similarly, there is a root between 0 and 1, and another between 1 and 2.

THEOREM 7. (Descartes' Rule of Signs.) *The number of positive real roots of an algebraic equation $f(x) = 0$ with real coefficients is either equal to the number of variations in sign of $f(x)$ or less than that number by a positive even integer. The number of negative real roots of $f(x) = 0$ is either equal to the number of variations in sign of $f(-x)$ or less than that number by a positive even integer.*

Example. $f(x) \equiv 8x^3 - 12x^2 - 2x + 3$ has two changes in sign, and therefore there are either two or no positive roots of $f(x) = 0$. Also, $f(-x) \equiv -8x^3 - 12x^2 + 2x + 3$ has only one change in sign, and $f(x)$ must have one negative root.

THEOREM 8. *Every algebraic equation of odd degree, with real coefficients, and leading coefficient positive has at least one real root whose sign is opposite to that of the constant term.*

Example. Since $f(x) \equiv 8x^3 - 12x^2 - 2x + 3 = 0$ is of odd degree and the constant term is positive, it follows that there must be at least one negative root.

THEOREM 9. *If an algebraic equation $f(x) = 0$ with real coefficients has a root $a + bi$, where $b \neq 0$, and a and b are real, it also has the root $a - bi$.*

Example. Thus, $x^3 - 1 = 0$ has the root $-\frac{1}{2} + \frac{1}{2}\sqrt{3}\,i$, and therefore it has the root $-\frac{1}{2} - \frac{1}{2}\sqrt{3}\,i$. This theorem states that imaginary roots always occur in pairs.

PROBLEMS

1. Find all the roots of the following equations:

   (*a*) $x^3 + 2x^2 - 4x - 8 = 0$;
   (*b*) $2x^3 - x^2 - 5x - 2 = 0$;
   (*c*) $4x^4 + 4x^3 + 3x^2 - x - 1 = 0$;
   (*d*) $2x^4 - 3x^3 - 3x - 2 = 0$.

2. Isolate the roots of the following equations between consecutive integers:

(a) $x^3 - 2x^2 - x + 1 = 0$;
(b) $2x^3 + 4x^2 - 2x - 3 = 0$;
(c) $x^3 + 5x^2 + 6x + 1 = 0$;
(d) $x^4 - 5x^2 + 3 = 0$.

28. Horner's Method. Many readers are already familiar with Horner's method of determining the value, to any desired number of decimal places, of the real roots of algebraic equations. However, the development given here is somewhat different from that used in the texts on algebra, in that it depends on Taylor's series expansion.

Suppose that the equation is

$$(28\text{-}1) \qquad f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0$$

and that it is known that the equation has a root between c and $c + 1$, where c is an integer. If $f(x)$ is expanded in Taylor's series in powers of $x - c$, there will result* a polynomial in $x - c$, namely,

$$f(x) = f(c) + f'(c)(x - c) + \frac{f''(c)}{2!}(x - c)^2 + \cdots + \frac{f^{(n-1)}(c)}{(n-1)!}(x - c)^{n-1} + \frac{f^{(n)}(c)}{n!}(x - c)^n.$$

Now, let $x - c = x_1$ and $\dfrac{f^{(r)}(c)}{r!} = A_{n-r}$. Then (28-1) is replaced by

$$(28\text{-}2) \quad f_1(x_1) \equiv A_n + A_{n-1}x_1 + \cdots + A_1x_1^{n-1} + A_0x_1^n = 0.$$

Since (28-1) had a root between c and $c + 1$ and since $x_1 = x - c$, it is evident that (28-2) has a root between 0 and 1. By the use of Theorem 6, Sec. 27, this root can be isolated between d and $d + 0.1$, where d has the form $a/10$ and $0 \leq a < 9$. Moreover, $f_1(x_1) = f(x_1 + c)$; and it follows that, if f_1 has a root between d and $d + 0.1$, then f has a root between $c + d$ and $c + d + 0.1$. It should be noted that c may be negative but that d will always be positive or zero.

The function $f_1(x_1)$ can be expanded in Taylor's series in powers of $x_1 - d$; and, if $x_2 = x_1 - d$, there will be obtained an equation

$$f_2(x_2) = B_n + B_{n-1}x_2 + \cdots + B_1x_2^{n-1} + B_0x_2^n = 0.$$

But $f_1(x_1) = 0$ had a root between d and $d + 0.1$; and since $x_2 = x_1 - d$, it follows that $f_2(x_2) = 0$ will have a root between 0 and 0.1.

This process can be continued as long as desired, each step determining another decimal place of the root of the original equation (28-1).

* Since $f(x)$ is a polynomial of the nth degree, the derivatives of order higher than n are all zero.

The solution of a specific equation may help to clarify the procedure. Let it be required to find the values of the real roots of the equation

$$F(x) \equiv x^4 + x^3 - 3x^2 - 6x - 3 = 0.$$

Since there is only one variation in sign, $F(x)$ has at most one positive root. $F(-x)$ has three variations, and so there will be at most three negative roots. The only possibilities for rational roots are ± 1 and ± 3. Since $F(-1) = 0$, it follows that $x = -1$ is a root. Moreover, if $F(x)$ is divided by $x + 1$, the quotient is $f(x) = x^3 - 3x - 3$. Hence, the remaining roots of $F(x) = 0$ are the three roots of

$$f(x) \equiv x^3 - 3x - 3 = 0.$$

It is easily checked that $f(x) = 0$ has no rational roots. Also, $\Delta = 108 - 243$, so that there is only one real root which, since $f(2) = -1$ and $f(3) = 15$, must lie between 2 and 3. Therefore, $f(x)$ will be expanded in powers of $x - 2$. Since

$$\begin{aligned} f(x) &= x^3 - 3x - 3, & f(2) &= -1, \\ f'(x) &= 3x^2 - 3, & f'(2) &= 9, \\ f''(x) &= 6x, & f''(2) &= 12, \\ f'''(x) &= 6, & f'''(2) &= 6, \end{aligned}$$

the expansion becomes

$$f(x) = -1 + 9(x - 2) + \frac{12}{2!}(x - 2)^2 + \frac{6}{3!}(x - 2)^3.$$

Replacing $x - 2$ by x_1 gives

$$f_1(x_1) \equiv -1 + 9x_1 + 6x_1^2 + x_1^3 = 0.$$

Since the real root of this equation lies between 0 and 1, the x_1^2 and x_1^3 terms do not contribute very much to the value of $f_1(x_1)$. Hence, a first approximation to the root can be obtained by setting $9x_1 - 1 = 0$. This gives $x_1 = 1/9 = 0.111 \cdots$, and suggests that the root probably lies between 0.1 and 0.2. It is easy to show that $f_1(0.1) = -0.039$ and $f_1(0.2) = 1.048$; there is thus a root between 0.1 and 0.2, and it is evidently closer to 0.1. Therefore, $f(x) = 0$ has a root between 2.1 and 2.2.

Expanding $f_1(x_1)$ in powers of $x_1 - 0.1$ gives

$$f_1(x_1) = -0.039 + 10.23(x_1 - 0.1) + \frac{12.6}{2!}(x_1 - 0.1)^2 + \frac{6}{3!}(x_1 - 0.1)^3,$$

and replacing $x_1 - 0.1$ by x_2 yields

$$f_2(x_2) = -0.039 + 10.23x_2 + 6.3x_2^2 + x_2^3 = 0.$$

Now $10.23x_2 - 0.039 = 0$ gives the approximation $x_2 = 0.0038$, and testing 0.003 and 0.004 reveals that $f_2(0.003) = -0.008253273$ and

$f_2(0.004) = +0.002020864$. Thus, the root lies between 0.003 and 0.004 and is closer to 0.004. If it is desired to determine the root of $f(x) = 0$ to three decimal places only, this value will be 2.104. If more decimal places are desired, the process can be continued. It should be noted that in each succeeding step the terms of the second and third degree contribute less, so that the linear approximation becomes better.

PROBLEMS

1. Apply Horner's method to find the cube root of 25, correct to three decimal places.

2. Determine the real roots of $x^3 - 2x - 1 = 0$ by Horner's method.

3. Determine the root of $x^4 + x^3 - 7x^2 - x + 5 = 0$, which lies between 2 and 3.

4. Determine the real root of $2x^3 - 3x^2 + x - 1 = 0$.

5. Determine the roots of $x^3 - 3x^2 + 3 = 0$.

6. Find, correct to three decimal places, the value of the root of $x^5 + 3x^3 - 2x^2 + x + 1 = 0$, which lies between -1 and 0.

7. A sphere 2 ft. in diameter is formed of wood whose specific gravity is $\frac{2}{3}$. Find to three significant figures the depth h to which the sphere will sink in water. $\left[\text{The volume of a spherical segment is } \pi h^2\left(r - \frac{h}{3}\right).\right]$ The volume of the submerged segment is equal to the volume of the displaced water, which must weigh as much as the sphere. Since water weighs 62.5 lb. per cubic foot,

$$\pi h^2\left(r - \frac{h}{3}\right) 62.5 = \frac{4}{3}\pi r^3 \frac{2}{3} 62.5$$

and, since $r = 1$,

$$h^3 - 3h^2 + \tfrac{8}{3} = 0.$$

29. Newton's Method. Horner's method of obtaining a numerical solution of an equation is probably the most useful scheme for solving algebraic equations, but it is not applicable to trigonometric, exponential, or logarithmic equations. A method applicable to these types as well as to algebraic equations was developed by Sir Isaac Newton sometime before 1676.

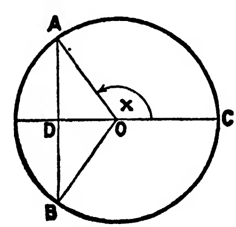

FIG. 19.

Newton applied his method to an algebraic equation, but it will be introduced here in the solution of a problem involving a trigonometric function.

Let it be required to find the angle subtended at the center of a circle by an arc whose length is double the length of its chord

(Fig. 19). Let the arc BCA be an arc of length $2BA$. Let $2x$ be the angle (measured in radians) subtended at the center of the circle. Then, arc $BCA = 2xr$ and $BA = 2\ DA = 2r \sin x$. If arc $BCA = 2BA$, then $2xr = 4r \sin x$, or $x - 2 \sin x = 0$.

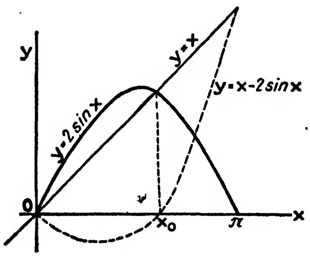

FIG. 20.

The graphical solution of equations of this type was discussed in Sec. 25. A first approximation can be obtained by graphical means. If $y = x$ and $y = 2 \sin x$ are plotted, it appears from the graph (Fig. 20) that they intersect for x lying between 108° and 109°, or, expressing this in radians,

$$1.8850 < x < 1.9024.$$

If $x_1 = 1.8850$ be chosen as the first approximation, the question of improving this value will be discussed first from the following graphical considerations.

If the part of the curve $y = x - 2 \sin x$ in the vicinity of the root be drawn on a large scale, it will have the appearance shown in Fig. 21. It is clear from the graph that adding to x_1 the

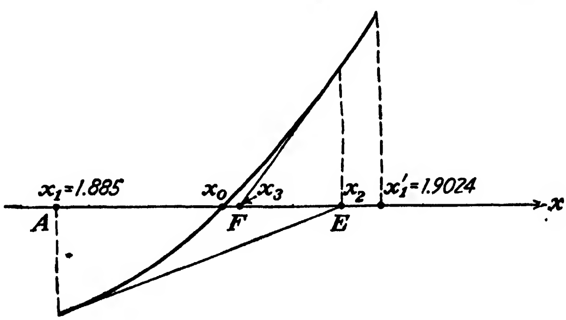

FIG. 21.

distance AE, cut off by the tangent line to the curve at $x_1 = 1.8850$, will give a value x_2 which is a better approximation to the actual root x_0. But AE is the subtangent at x_1 and is equal to $-\dfrac{f(x_1)}{f'(x_1)}$, where $f(x) = x - 2 \sin x$. Thus,*

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

* See, in this connection, Prob. 8, at the end of this section.

Similarly, upon using x_2 as the second approximation and observing that $-\dfrac{f(x_2)}{f'(x_2)}$ is the subtangent EF, the third approximation is found to be

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)},$$

and in general the nth approximation x_n is given by

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, \quad (n = 2, 3, \cdots). \tag{29-1}$$

Since $x_1 = 1.885$, the formula gives

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = x_1 - \frac{x_1 - 2\sin x_1}{1 - 2\cos x_1}$$
$$= 1.8850 - \frac{1.8850 - 1.9022}{1 + 0.6180} = 1.8956.$$

In a similar way,

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 1.8956 - \frac{1.8956 - 2\sin 1.8956}{1 - 2\cos 1.8956} = 1.8955.$$

It follows that the angle subtended by the arc is 3.7910 radians.

The use of Newton's method requires some preliminary examination of the equation. It may happen that the equation is of such a character that the second approximation to x_0 will be worse than the first. A careful examination of the following sketches of four types of functions, sketched in the vicinity of their roots, reveals the fact that some care must be exercised in applying Newton's method. For all four figures, it is assumed that x_0 has been isolated between x_1 and x_1'. The graphical interpretation of the correction $-\dfrac{f(x_1)}{f'(x_1)}$ as the subtangent must be kept in mind throughout this discussion. If x_1 is used as the first approximation, then x_2 will be obtained as the second approximation by using Newton's method; if x_1' is used, then x_2' will be obtained.

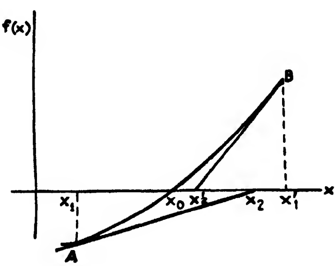

FIG. 22.

In Fig. 22, both x_2 and x_2' are closer to x_0 than x_1 or x_1'. In this case the method would work regardless of which value is chosen

as the first approximation. In Fig. 23, x_2 is better than the first approximation x_1, but x_2' is worse than x_1'. It appears from

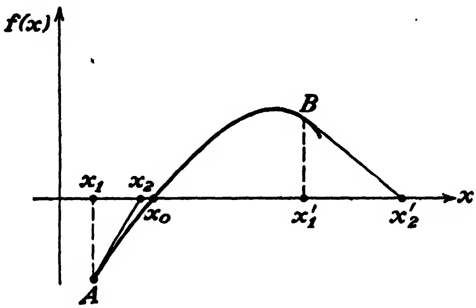

Fig. 23.

the figure that this occurs because the curve is concave down between x_1 and x_1', and hence $f''(x) < 0$, whereas $f(x_1) < 0$ and $f(x_1') > 0$. A similar situation would obtain if the curve is concave up, so that $f''(x) > 0$ (Fig. 24). The reader will readily convince himself from an inspection of Fig. 23 that caution must be exercised in the choice of the first approximation if the curve has a maximum (or a minimum) in the vicinity of x_0.

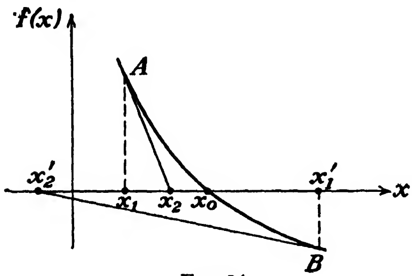

Fig. 24.

If the curve has the appearance indicated in Fig. 25, then it is evident that the choice of either x_1 or x_1' as the first approximation will yield a second approximation which is worse than the first one. This is due to the fact that the curve has a point of inflection between x_1 and x_1'.

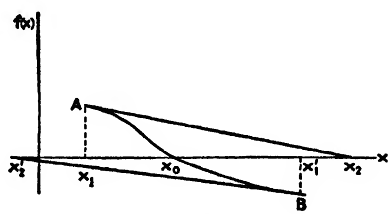

Fig. 25.

From the foregoing discussion, it is apparent that Newton's method should not be applied before making an investigation of the behavior of the first and second derivatives of $f(x)$ in the vicinity of the root. The

conclusions drawn from this discussion can be summarized in the following practical rule for determining the choice of the first approximation: *If $f'(x)$ and $f''(x)$ do not vanish in the given interval (x_1, x_1') and if the signs of $f(x_1)$ and $f(x_1')$ are opposite, then the first approximation should be chosen as that one of the two end points for which $f(x)$ and $f''(x)$ have the same sign.*

It can be proved* that if the single-valued continuous function $f(x)$ is of such a nature that $f(x) = 0$ has only one real root in (x_1, x_1') and both $f'(x)$ and $f''(x)$ are continuous and do not vanish in (x_1, x_1'), then repeated applications of Newton's method will determine the value of the root of $f(x) = 0$ to any desired number of decimal places.

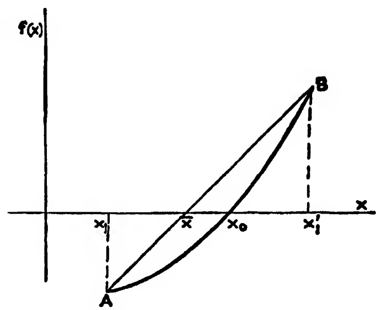

FIG. 26.

The cases to which Newton's method does not apply can be treated by a method of interpolation (*regula falsi*) that is applicable to any equation.

Let $\bar{x}$ be the value of x for which the chord AB intersects the x-axis. From similar triangles (Fig. 26),

$$\frac{\bar{x} - x_1}{-f(x_1)} = \frac{x_1' - \bar{x}}{f(x_1')}.$$

Solving for $\bar{x}$ gives

$$\bar{x} = \frac{x_1 f(x_1') - x_1' f(x_1)}{f(x_1') - f(x_1)}.$$

The value $\bar{x}$ is clearly a better approximation than either x_1 or x_1'.

PROBLEMS

1. Solve Prob. 7, Sec. 28, by Newton's method. Also, apply the method of interpolation.

2. Determine the angle subtended at the center of a circle by a chord which cuts off a segment whose area is one-quarter of that of the circle.

3. Find the roots of $e^x - 4x = 0$, correct to four decimal places.

4. Solve $x - \cos x = 0$.

* See WEBER, H., Algebra, 2d ed. vol. 1, pp. 380–382; COATE, G. T., On the Convergence of Newton's Method of Approximation, *Amer. Math. Monthly*, vol. 44, pp. 464–466, 1937.

5. Solve $x = \tan x$ in the vicinity of $x = \frac{3}{2}\pi$.
6. Solve $x + e^x = 0$.
7. Solve $x^4 - x - 1 = 0$.
8. Show with the aid of Taylor's series that, if $x = x_1$ is an approximate root of $f(x) = 0$, then the nth approximation is, in general, determined from the formula (29-1).

Hint: $f(x) = f(x_1) + f'(x_1)(x - x_1) + \cdots + \frac{f^{(n)}(x_1)}{n!}(x - x_1)^n + \cdots$;
and if $f(x_2) \doteq 0$, then

$$0 \doteq f(x_1) + f'(x_1)(x_2 - x_1).$$

30. Determinants of the Second and Third Order. The solution of systems of linear equations involves the determination of the particular values of two or more variables that will satisfy simultaneously a set of equations in those variables. Since the discussion is simplified by using certain properties of determinants and matrices, the remainder of this chapter is devoted to some elementary theory of determinants and matrices and its application to the solution of systems of linear equations.

Consider first a system composed of two linear equations in two unknowns, namely,

$$\text{(30-1)} \qquad \begin{cases} a_1x + b_1y = k_1, \\ a_2x + b_2y = k_2. \end{cases}$$

If y is eliminated between these two equations, there is obtained the equation

$$\text{(30-2)} \qquad (a_1b_2 - a_2b_1)x = k_1b_2 - k_2b_1;$$

and if x is eliminated, there results

$$\text{(30-3)} \qquad (a_1b_2 - a_2b_1)y = a_1k_2 - a_2k_1.$$

If the expression $a_1b_2 - a_2b_1$ is not zero, the two equations (30-2) and (30-3) can be solved to give values for x and y. That the values so obtained are actually the solutions of the system (30-1) can be verified by substitution in Eqs. (30-1).

The expression $a_1b_2 - a_2b_1$ occurs as the coefficient for both x and y. Denote it by the symbol

$$\text{(30-4)} \qquad a_1b_2 - a_2b_1 \equiv \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

This symbol is called a determinant of the second order. It is also called the determinant of the coefficients of the system

(30-1), for the elements of its first column are the coefficients of x and the elements of its second column are the coefficients of y. Using this notation, (30-2) and (30-3) become

$$\text{(30-5)} \qquad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} x = \begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix}, \qquad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} y = \begin{vmatrix} a_1 & k_1 \\ a_2 & k_2 \end{vmatrix}.$$

The definition (30-4) provides the method of evaluating the symbol. If

$$D \equiv \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0,$$

the unique solution of (30-1) can be written as

$$x = \frac{\begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix}}{D}, \qquad y = \frac{\begin{vmatrix} a_1 & k_1 \\ a_2 & k_2 \end{vmatrix}}{D}.$$

If $D = 0$, $a_1b_2 = a_2b_1$ or $a_1/a_2 = b_1/b_2$. But if the corresponding coefficients of the two equations are proportional, the two lines, whose equations are given by (30-1), are parallel (if $a_1/a_2 \neq k_1/k_2$) or coincident (if $a_1/a_2 = b_1/b_2 = k_1/k_2$). In the first case, the determinants appearing as the right-hand members of the equations in (30-5) will be different from zero and there will be no solution for x and y. In the second case, these determinants, as well as D, are zero and any pair of values x, y that satisfies one equation of the system will satisfy the other equation, also.

Example 1. For the system

$$\begin{aligned} 2x - 3y &= -4 \\ 3x - y &= 1, \end{aligned} \qquad D = \begin{vmatrix} 2 & -3 \\ 3 & -1 \end{vmatrix} = -2 + 9 = 7,$$

$$x = \frac{\begin{vmatrix} -4 & -3 \\ 1 & -1 \end{vmatrix}}{7} = 1, \qquad y = \frac{\begin{vmatrix} 2 & -4 \\ 3 & 1 \end{vmatrix}}{7} = 2.$$

Example 2. For the system

$$\begin{aligned} 2x - 3y &= 4 \\ 6x - 9y &= 5, \end{aligned} \qquad D = \begin{vmatrix} 2 & -3 \\ 6 & -9 \end{vmatrix} = 0,$$

but

$$\frac{2}{6} = \frac{-3}{-9} \neq \frac{4}{5}.$$

The two lines whose equations are given are parallel.

Example 3. For the system

$$\begin{aligned} 2x - 3y &= 4 \\ 6x - 9y &= 12, \end{aligned} \qquad D = \begin{vmatrix} 2 & -3 \\ 6 & -9 \end{vmatrix} = 0,$$

$$\frac{2}{6} = \frac{-3}{-9} = \frac{4}{12}.$$

The two lines are coincident.

Consider next the system of three linear equations in three unknowns,

$$\text{(30-6)} \qquad \begin{cases} a_1x + b_1y + c_1z = k_1, \\ a_2x + b_2y + c_2z = k_2, \\ a_3x + b_3y + c_3z = k_3. \end{cases}$$

If these equations are multiplied, respectively, by

$$b_2c_3 - b_3c_2, \qquad b_3c_1 - b_1c_3, \qquad b_1c_2 - b_2c_1,$$

and the resulting equations are added, the sum is

$$\text{(30-7)} \qquad \begin{aligned} (a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1)x \\ = k_1b_2c_3 - k_1b_3c_2 + k_2b_3c_1 - k_2b_1c_3 + k_3b_1c_2 - k_3b_2c_1. \end{aligned}$$

The coefficient of x in (30-7) can be denoted by the symbol

$$\text{(30-8)} \qquad D \equiv \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \equiv a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1.$$

This symbol is called a determinant of the third order. It is also the determinant of the coefficients of the system (30-6).

Using the notation of (30-8), Eq. (30-7) can be written as

$$Dx \equiv \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} x = \begin{vmatrix} k_1 & b_1 & c_1 \\ k_2 & b_2 & c_2 \\ k_3 & b_3 & c_3 \end{vmatrix}.$$

Similarly it can be shown that

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} y = \begin{vmatrix} a_1 & k_1 & c_1 \\ a_2 & k_2 & c_2 \\ a_3 & k_3 & c_3 \end{vmatrix},$$

and

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} z = \begin{vmatrix} a_1 & b_1 & k_1 \\ a_2 & b_2 & k_2 \\ a_3 & b_3 & k_3 \end{vmatrix}.$$

If $D \neq 0$, the unique solutions for x, y, and z can be obtained as

$$(30\text{-}9)\qquad x = \frac{\begin{vmatrix} k_1 & b_1 & c_1 \\ k_2 & b_2 & c_2 \\ k_3 & b_3 & c_3 \end{vmatrix}}{D}, \qquad y = \frac{\begin{vmatrix} a_1 & k_1 & c_1 \\ a_2 & k_2 & c_2 \\ a_3 & k_3 & c_3 \end{vmatrix}}{D}, \qquad z = \frac{\begin{vmatrix} a_1 & b_1 & k_1 \\ a_2 & b_2 & k_2 \\ a_3 & b_3 & k_3 \end{vmatrix}}{D}.$$

In order to show that the values of x, y, and z, given in (30-9), actually satisfy Eqs. (30-6), these values can be substituted in the given equations.

If $D = 0$, the three equations (30-6) are either inconsistent or dependent. A detailed analytic discussion of these cases will be given in Sec. 35. Since the three equations of (30-6) are the equations of three planes, a geometrical interpretation will now be given.

If the three equations are inconsistent, the three planes are all parallel, or two are parallel and are cut by the third plane in two parallel lines. In either case, there is obviously no solution for x, y, and z. If the equations are dependent, all three planes intersect in the same line or all three planes coincide. In either case there will be an infinite number of solutions for x, y, and z.

Example. For the system

$$\begin{aligned} 3x - y - z &= 2, \\ x - 2y - 3z &= 0, \\ 4x + y + 2z &= 4. \end{aligned} \qquad D = \begin{vmatrix} 3 & -1 & -1 \\ 1 & -2 & -3 \\ 4 & 1 & 2 \end{vmatrix} = 2.$$

Therefore,

$$x = \frac{\begin{vmatrix} 2 & -1 & -1 \\ 0 & -2 & -3 \\ 4 & 1 & 2 \end{vmatrix}}{2} = \frac{2}{2} = 1, \qquad y = \frac{\begin{vmatrix} 3 & 2 & -1 \\ 1 & 0 & -3 \\ 4 & 4 & 2 \end{vmatrix}}{2} = \frac{4}{2} = 2,$$

$$z = \frac{\begin{vmatrix} 3 & -1 & 2 \\ 1 & -2 & 0 \\ 4 & 1 & 4 \end{vmatrix}}{2} = -\frac{2}{2} = -1.$$

PROBLEMS

1. Evaluate

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & -1 & -2 \end{vmatrix}, \quad \begin{vmatrix} 2 & 0 & -3 \\ 1 & 4 & 2 \\ -1 & 1 & -2 \end{vmatrix}, \text{ and } \begin{vmatrix} 4 & -2 & 1 \\ 5 & 0 & -1 \\ 2 & 3 & -3 \end{vmatrix}.$$

2. Find the solutions of the following systems of equations by using determinants:

(a) $5x - 4y = 3,$
$2x + 3y = 7;$

(b) $2x + 3y - 2z = 4,$
$x + y - z = 2,$
$3x - 5y + 3z = 0;$

(c) $3x - 2y = 7,$
$3y + 2z = 6,$
$2x + 3z = 1;$

(d) $3x + 2y + 2z = 3,$
$x - 4y + 2z = 4,$
$2x + y + z = 2.$

31. Determinants of the *n*th Order. Determinants of the second and third orders were defined in the preceding section. These are merely special cases of the definition of the determinants of any order n. Instead of a symbol with 2^2 or 3^2 elements, the determinant of the nth order is defined as the symbol, with n rows and n columns,

$$D \equiv \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

which stands for the sum* of the $n!$ terms $(-1)^k a_{k_1 1} a_{k_2 2} \cdots a_{k_n n}$, where $k_1, k_2, \cdots, k_n$ are the numbers $1, 2, \cdots, n$ in some order. The integer k is defined as the number of *inversions of order* of the subscripts $k_1, k_2, \cdots, k_n$ from the normal order $1, 2, \cdots, n$, where a particular arrangement is said to have k inversions of order if it is necessary to make k successive interchanges of adjacent elements† in order to make the arrangement assume the normal order. There are $n!$ terms since there are $n!$ permutations of the n first subscripts. Moreover, it is evident that each term contains as a factor one and only one element from each row and one and only one element from each column.

* This sum is sometimes called the expansion of the determinant.

† It should be noted that it is not necessary to specify that the interchanges should be of adjacent elements, for it can be proved that, if any particular arrangement can be obtained by k interchanges of adjacent elements and also by k' interchanges of some other type, then k and k' are always either both even or both odd. Hence, the sign of the term is independent of the particular succession of interchanges.

Example. Consider the third-order determinant

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

The six terms of the expansion are, apart from sign,

$$\begin{matrix} a_{11}a_{22}a_{33}, & a_{11}a_{32}a_{23}, & a_{21}a_{12}a_{33}, \\ a_{21}a_{32}a_{13}, & a_{31}a_{12}a_{23}, & a_{31}a_{22}a_{13}. \end{matrix}$$

The first term, in which the first subscripts have the normal order, is called the diagonal term, and its sign is positive. In the second term the arrangement 132 requires the interchange of 2 and 3 to make it assume the normal order; therefore, $k = 1$, and the term has a negative sign. Similarly, the third term has a negative sign. The fourth term will have a positive sign, for the arrangement 231 requires the interchange of 3 and 1 followed by the interchange of 2 and 1 in order to assume the normal order. Similarly, it appears that the fifth term will have a positive sign. In the sixth term, it is necessary to make three interchanges (3 and 2, 3 and 1, and 2 and 1) in order to arrive at the normal order; hence, this term will have a negative sign. As a result of this investigation, it follows that

$$D = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}$$

It is evident that if k is equal to zero or an even number the term will have a positive sign, whereas if k is odd the term will be negative.

PROBLEM

Find the signs of the six terms involving a_{11} in the expansion of the determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}.$$

32. Properties of Determinants. 1. *The value of a determinant is not changed if in the symbol the elements of corresponding rows and columns are interchanged.*

If

$$D \equiv \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

then the determinant formed by interchanging the corresponding rows and columns is

$$D' \equiv \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix}.$$

Any term $(-1)^k a_{k_1 1} a_{k_2 2} \cdots a_{k_n n}$ of D, where $k_1, k_2, \cdots, k_n$ are the numbers $1, 2, \cdots, n$ in some order, will correspond to a term $(-1)^l a_{k_1 1} a_{k_2 2} \cdots a_{k_n n}$ of D', for each determinant must contain every possible term that is a product of one and only one element from each row and each column. But the number of inversions is the same for the term of D as it is for the term of D', owing to the fact that the corresponding first subscripts are the same. It follows that each term of D occurs also in D', and conversely each term of D' occurs also in D.

Example. If

$$D \equiv \begin{vmatrix} 2 & 5 & 3 \\ 1 & -1 & 4 \\ -3 & -2 & 1 \end{vmatrix} = -66,$$

then

$$D' \equiv \begin{vmatrix} 2 & 1 & -3 \\ 5 & -1 & -2 \\ 3 & 4 & 1 \end{vmatrix} = -66.$$

2. *An interchange of any two rows or of any two columns of a determinant will merely change the sign of the determinant.*

If D is the original determinant and D'' is the determinant having the ith and jth rows of D interchanged, then the expansion of D'' will have the first subscripts of each term the same as those of the corresponding term of D, except that i and j will be interchanged. Since it requires one interchange to restore i and j to their original order in each term, the sign of every term will be changed. Thus, $D'' = -D$.

Example. If

$$D \equiv \begin{vmatrix} 2 & 5 & 3 \\ 1 & -1 & 4 \\ -3 & -2 & 1 \end{vmatrix} = -66,$$

then

$$D'' = \begin{vmatrix} 2 & 5 & 3 \\ -3 & -2 & 1 \\ 1 & -1 & 4 \end{vmatrix} = 66.$$

3. *If any two rows or any two columns of a determinant are identical, the value of the determinant is zero.*

For, by property 2, if these two rows (or columns) were interchanged, the sign of D should be changed. But since these two rows (or columns) are identical, D remains unchanged. Therefore, $D = -D$, and hence $D = 0$.

Example. If

$$D \equiv \begin{vmatrix} 2 & -1 & 2 \\ 3 & 4 & 3 \\ -2 & 5 & -2 \end{vmatrix},$$

then

$$D = 0.$$

4. *If each element of any row or any column be multiplied by m, the value of the determinant is multiplied by m.*

This follows from the definition of the determinant. Since one and only one element of any row or column occurs in each term, each term will be multiplied by m and therefore the value of the determinant is multiplied by m.

Example 1. If

$$D \equiv \begin{vmatrix} 2 & 5 & 3 \\ 1 & -1 & 4 \\ -3 & -2 & 1 \end{vmatrix} = -66$$

and

$$\bar{D} \equiv \begin{vmatrix} 2 & 5 & 3 \\ 1 & -1 & 4 \\ -6 & -4 & 2 \end{vmatrix},$$

which has each element of the last row twice the corresponding element of the last row of D, then

$$\bar{D} = -132 \qquad \text{and} \qquad \bar{D} = 2D.$$

Example 2. If

$$D \equiv \begin{vmatrix} 6 & 4 & 8 \\ 9 & 2 & 1 \\ -6 & 3 & -1 \end{vmatrix},$$

then

$$D \equiv 2\begin{vmatrix} 3 & 2 & 4 \\ 9 & 2 & 1 \\ -6 & 3 & -1 \end{vmatrix} = 2 \cdot 3\begin{vmatrix} 1 & 2 & 4 \\ 3 & 2 & 1 \\ -2 & 3 & -1 \end{vmatrix}.$$

5. *From properties 3 and 4, it follows that the value of a determinant is zero if any two rows or any two columns have corresponding elements proportional.*

6. *The product of two determinants D and D', both of order n, is the nth-order determinant D'' which has as the element in its ith row and jth column the sum*

$$\sum_{k=1}^{n} a_{ik}b_{kj} \equiv a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj},$$

which is formed by multiplying each element a_{ik} of the ith row of D by the corresponding element b_{kj} of the jth column of D'.

Thus, if

$$D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \text{and} \quad D' = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix},$$

then

$$D \cdot D' \equiv D'' = \begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix}.$$

Example. The product of the following determinants is easily found by expanding the product determinant:

$$\begin{vmatrix} \sin x & \cos x & 1 \\ \sec x & \tan x & 1 \\ \csc x & \cot x & 1 \end{vmatrix} \cdot \begin{vmatrix} \sin x & \sec x & \csc x \\ \cos x & -\tan x & -\cot x \\ -1 & -1 & -1 \end{vmatrix}$$

$$= \begin{vmatrix} 0 & \tan x - \sin x - 1 & -\cos x \cot x \\ \tan x + \sin x - 1 & 0 & \sec x \csc x - 2 \\ \cos x \cot x & \sec x \csc x - 2 & 0 \end{vmatrix}$$

$$= 2\cos^2 x(2 - \sec x \csc x).$$

33. Minors. The method of evaluating a determinant by the use of the definition of Sec. 31 is exceedingly tedious, especially if $n \geq 4$. There are other schemes for this evaluation, and these require the definition of the minors of a determinant. The simplest of these schemes will be described and used here.

If, in the determinant D, the ith row and the jth column be suppressed, the resulting determinant A_{ij} (which is of order one less than the order of D) is called the minor of the element a_{ij}, which is in the ith row and jth column.

Example. If

$$D \equiv \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix},$$

then

$$A_{23} \equiv \begin{vmatrix} a_{11} & a_{12} & a_{14} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix}.$$

From the definition of a determinant, it is evident that $a_{ij}A_{ij}$ is composed of all the terms of D which contain the element a_{ij} as a factor, except for the possibility that all the signs may be reversed. Then the expression $(-1)^{k_1}a_{11}A_{11}$ is composed of all the terms of D containing a_{11} as a factor; $(-1)^{k_2}a_{21}A_{21}$ is composed of all the terms containing a_{21} as a factor; $(-1)^{k_3}a_{31}A_{31}$ is composed of all the terms containing a_{31} as a factor; etc. But D is composed of all the terms containing $a_{11}, a_{21}, a_{31}, \cdots, a_{n1}$ as a factor, and so

$$D = (-1)^{k_1}a_{11}A_{11} + (-1)^{k_2}a_{21}A_{21} + \cdots + (-1)^{k_n}a_{n1}A_{n1}.$$

It can be proved* that $k_1 = 1 + 1$. $k_2 = 2 + 1$, $k_3 = 3 + 1$, $\cdots$, $k_n = n + 1$, so that

$$D = a_{11}A_{11} - a_{21}A_{21} + \cdots + (-1)^{n+1}a_{n1}A_{n1}.$$

In the above development for D the elements $a_{11}, a_{21}, \cdots, a_{n1}$ are the elements of the first column of D. Similarly, the value of D can be formed by taking the elements of any other column or of any row.

Using the ith column gives

$$D = (-1)^{k_1}a_{1i}A_{1i} + (-1)^{k_2}a_{2i}A_{2i} + \cdots + (-1)^{k_n}a_{ni}A_{ni},$$

where $k_1 = i + 1$, $k_2 = i + 2$, $\cdots$, $k_n = i + n$. Similarly, using the ith row gives

$$D = (-1)^{k_1}a_{i1}A_{i1} + (-1)^{k_2}a_{i2}A_{i2} + \cdots + (-1)^{k_n}a_{in}A_{in},$$

where $k_1 = i + 1$, $k_2 = i + 2$, $\cdots$, $k_n = i + n$. It may be observed that each k_r is equal to the sum of the subscripts of its a_{ij} and is thus equal to the sum of the number of the row and the number of the column in which this element occurs. This development is known as the *expansion by minors*, or the *simple Laplace development.*

Since the term *cofactor* is frequently used in applications of this type of development, it will be defined here. *The cofactor C_{ij} of an element a_{ij} is defined as the signed minor, that is,*

$$C_{ij} = (-1)^{i+j}A_{ij}.$$

* DICKSON, L. E., First Course in Theory of Equations, pp. 101–127; FINE, H. B., College Algebra, pp. 492–519.

Thus, the expression for D can be written as

$$D = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} A_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} A_{ij},$$

or as

$$D = \sum_{i=1}^{n} a_{ij} C_{ij} = \sum_{j=1}^{n} a_{ij} C_{ij}.$$

Example.

$$\begin{vmatrix} 3 & 4 & 0 & 6 \\ 0 & 5 & 2 & 1 \\ 0 & 3 & 4 & 0 \\ 1 & 2 & 7 & 1 \end{vmatrix} = 3\begin{vmatrix} 5 & 2 & 1 \\ 3 & 4 & 0 \\ 2 & 7 & 1 \end{vmatrix} - 0\begin{vmatrix} 4 & 0 & 6 \\ 3 & 4 & 0 \\ 2 & 7 & 1 \end{vmatrix} + 0\begin{vmatrix} 4 & 0 & 6 \\ 5 & 2 & 1 \\ 2 & 7 & 1 \end{vmatrix} - 1\begin{vmatrix} 4 & 0 & 6 \\ 5 & 2 & 1 \\ 3 & 4 & 0 \end{vmatrix}$$

$$= 3\left[-3\begin{vmatrix} 2 & 1 \\ 7 & 1 \end{vmatrix} + 4\begin{vmatrix} 5 & 1 \\ 2 & 1 \end{vmatrix} - 0\begin{vmatrix} 5 & 2 \\ 2 & 7 \end{vmatrix}\right]$$

$$- 1\left[4\begin{vmatrix} 2 & 1 \\ 4 & 0 \end{vmatrix} - 0\begin{vmatrix} 5 & 1 \\ 3 & 0 \end{vmatrix} + 6\begin{vmatrix} 5 & 2 \\ 3 & 4 \end{vmatrix}\right]$$

$$= -9(2 - 7) + 12(5 - 2) - 4(0 - 4) - 6(20 - 6)$$

$$= 13.$$

Here, the first expansion is made by using the elements of the first column, for it contains two zeros (the third row is an equally good choice). The expansion of the first third-order determinant is made by using the elements of the second row, but the third column could be used to equal advantage. In the expansion of the last third-order determinant the first row was chosen, but the third row and the second and third columns provide equally good choices.

The following theorem is given here because of its frequent use in many fields of pure and applied mathematics:

THEOREM. *The sum* $\sum_{j=1}^{n} a_{ij} C_{kj}$ *is zero*, if $k \neq i$.

Each term of this sum is formed by taking the product of the cofactor of an element of the kth row by the corresponding element of the ith row. This is the expansion of a determinant whose ith and kth rows are identical and whose value is accordingly zero. Similarly, it follows that $\sum_{i=1}^{n} a_{ij} C_{ik} = 0$, if $k \neq j$.

Example. Let

$$D \equiv \begin{vmatrix} 3 & -1 & 2 \\ 1 & 2 & -1 \\ 4 & -3 & -2 \end{vmatrix}.$$

Then,

$$C_{11} = -7, \qquad C_{12} = -2, \qquad C_{13} = -11$$

and the sum

$$\sum_{j=1}^{3} a_{3j}C_{1j} = -28 + 6 + 22 = 0.$$

Similarly,

$$\sum_{j=1}^{3} a_{2j}C_{1j} = -7 - 4 + 11 = 0.$$

By using the theory of determinants, the solution of a system of n non-homogeneous linear equations in n unknowns can be obtained. The rule for effecting the solution will be stated but not proved.* The proof for the cases when $n = 2$ and $n = 3$ has already been given in Sec. 30.

Cramer's Rule. Let

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= k_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= k_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= k_n \end{aligned} \tag{33-1}$$

be a system of n equations in the unknowns $x_1, x_2, \cdots, x_n$, such that the determinant

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

of the coefficients is not zero. The system (33-1) has a unique solution given by

$$x_1 = \frac{D_1}{D}, \qquad x_2 = \frac{D_2}{D}, \quad \cdots, \qquad x_n = \frac{D_n}{D},$$

where D_i is the determinant formed by replacing the elements $a_{1i}, a_{2i}, a_{3i}, \cdots, a_{ni}$ of the ith column of D by $k_1, k_2, k_3, \cdots, k_n$, respectively.

Example. Solve, by Cramer's rule, the system

$$\begin{aligned} 3x + y + 2z &= 3, \\ 2x - 3y - z &= -3, \\ x + 2y + z &= 4. \end{aligned}$$

* Dickson, L. E., First Course in Theory of Equations, pp. 114–115.

Here

$$D = \begin{vmatrix} 3 & 1 & 2 \\ 2 & -3 & -1 \\ 1 & 2 & 1 \end{vmatrix} = 8$$

and

$$x = \frac{\begin{vmatrix} 3 & 1 & 2 \\ -3 & -3 & -1 \\ 4 & 2 & 1 \end{vmatrix}}{8} = \frac{8}{8} = 1, \qquad y = \frac{\begin{vmatrix} 3 & 3 & 2 \\ 2 & -3 & -1 \\ 1 & 4 & 1 \end{vmatrix}}{8} = \frac{16}{8} = 2,$$

$$z = \frac{\begin{vmatrix} 3 & 1 & 3 \\ 2 & -3 & -3 \\ 1 & 2 & 4 \end{vmatrix}}{8} = \frac{-8}{8} = -1.$$

PROBLEMS

1. Evaluate

$$\begin{vmatrix} 1 & -2 & 0 & 3 \\ -1 & 4 & 1 & 2 \\ 2 & 0 & -1 & -3 \\ 3 & 1 & 0 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2 & -1 & 3 & 0 \\ 1 & 4 & 0 & 2 \\ 1 & 1 & 3 & 5 \\ -1 & 0 & 2 & 0 \end{vmatrix}, \text{ and } \begin{vmatrix} 1 & 3 & 2 & -1 \\ 0 & 4 & -3 & 2 \\ -3 & 1 & 0 & 1 \\ 1 & 2 & 0 & -4 \end{vmatrix}.$$

2. Solve, by Cramer's rule, the systems

(a) $x + 2y + 3z = 3,\quad 2x - y + z = 6,\quad 3x + y - z = 4.$

(b) $2x + y + 3z = 2,\quad 3x - 2y - 2z = 1,\quad x - y + z = -1.$

(c) $x + 2y = 1,\quad 2x - y - 2z = 3,\quad -x + y + 3z = 2.$

(d) $2x + y + 3z + w = -2,\quad 5x + 3y - z - w = 1,\quad x - 2y + 4z + 3w = 4,\quad 3x - y + z = 2.$

34. Matrices and Linear Dependence. In order to discuss the systems arising in the succeeding sections, it is convenient to give a short introduction to the theory of matrices.*

An $m \times n$ matrix is defined as a system of mn quantities a_{ij} arranged in a rectangular array of m rows and n columns. If $m = n$, the array is called a square matrix of order n. The quantities a_{ij} are called the elements of the matrix. Thus,

$$(34\text{-}1) \qquad A \equiv \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{Vmatrix} \quad \text{or} \quad \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

* For detailed treatment see M. Bocher, Introduction to Higher Algebra, pp. 20–53; L. E. Dickson, Modern Algebraic Theories, pp. 39–63.

where double bars or parentheses are used to enclose the array of elements. If the order of the elements in (34-1) is changed or if any element is changed, a different matrix results. Any two matrices A and B are said to be equal if and only if every element of A is equal to the corresponding element of B, that is, if $a_{ij} = b_{ij}$ for every i and j.

If the matrix is square, it is possible to form from the elements of the matrix a determinant whose elements have the same arrangement as those of the matrix. The determinant is called the *determinant of the matrix*. From any matrix, other matrices can be obtained by striking out any number of rows and columns. Certain of these matrices will be square matrices, and the determinants of these matrices are called the determinants of the matrix. For an $m \times n$ matrix, there are square matrices of orders $1, 2, \cdots, p$, where p is equal to the smaller of the numbers m and n.

Example. The 2×3 matrix

$$A \equiv \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

contains the first-order square matrices (a_{11}), (a_{12}), (a_{23}), etc., obtained by striking out any two columns and any one row. It also contains the second-order square matrices

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix}, \quad \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix},$$

obtained by striking out any column of A.

In many applications, it is useful to employ the notion of the rank of a matrix A. This is defined in terms of the determinants of A. *A matrix A is said to be of rank r if there exists at least one r-rowed determinant of A that is not zero, whereas all determinants of A of order higher than r are zero.**

Example. If

$$A \equiv \begin{pmatrix} 1 & 0 & 1 & 3 \\ 2 & 1 & 0 & -2 \\ -1 & -1 & 1 & 5 \end{pmatrix},$$

* In case an $m \times n$ matrix contains no determinants of order higher than r, obviously r is the smaller of the numbers m and n, and the matrix is said to be of rank r.

the third-order determinants are

$$\begin{vmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & 3 \\ 2 & 1 & -2 \\ -1 & -1 & 5 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 1 & 3 \\ 2 & 0 & -2 \\ -1 & 1 & 5 \end{vmatrix} = 0,$$

$$\begin{vmatrix} 0 & 1 & 3 \\ 1 & 0 & -2 \\ -1 & 1 & 5 \end{vmatrix} = 0.$$

Since

$$\begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} \neq 0,$$

there is at least one second-order determinant different from zero, whereas all third-order determinants of A are zero. Therefore, the rank of A is 2.

It should be observed that a matrix is said to have *rank zero if all of its elements are zero.*

The notion of linear dependence is of importance in connection with the study of systems of linear equations, and it will be considered next.

A set of m, $m \geq 2$, quantities $f_1, f_2, f_3, \cdots, f_m$ (which may be constants or functions of any number of variables) is said to be linearly dependent if there exist m constants $c_1, c_2, \cdots, c_m$, which are not all zero, such that

$$c_1 f_1 + c_2 f_2 + \cdots + c_m f_m \equiv 0. \tag{34-2}$$

If no such constants exist, the quantities f_i are said to be *linearly independent.*

Example. If the f_i are the polynomials

$$\begin{aligned} f_1(x, y, z) &\equiv 2x^2 - 3xy + 4z, \\ f_2(x, y, z) &\equiv x^2 + 2xy - 3z, \\ f_3(x, y, z) &\equiv 4x^2 + xy - 2z, \end{aligned}$$

and if the constants are chosen as $c_1 = 1$, $c_2 = 2$, $c_3 = -1$, then

$$c_1 f_1 + c_2 f_2 + c_3 f_3 \equiv 0.$$

Therefore, these three polynomials are linearly dependent.

It is evident that, whenever the set of quantities is linearly dependent, at least one of the f_i can be expressed as a linear combination of the others. Thus, from (34-2), if $c_1 \neq 0$, then

$$f_1 = a_2 f_2 + a_3 f_3 + \cdots + a_m f_m,$$

where

$$a_2 = -\frac{c_2}{c_1}, \qquad a_3 = -\frac{c_3}{c_1}, \text{ etc.}$$

The definition of linear dependence requires the existence of at least one constant $c_i \neq 0$, and therefore the solution for f_i is assured.

Obviously, in most cases it would be extremely difficult to apply the definition in order to establish the linear dependence (or independence) of a given set of quantities. In case the quantities f_i are linear functions of n variables, there is a simple test which will be stated without proof.*

THEOREM. *The m linear functions*

$$f_i \equiv a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n, \quad (i = 1, 2, \cdots, m),$$

are linearly dependent if and only if the matrix of the coefficients is of rank $r < m$. Moreover, there are exactly r of the f_i that form a linearly independent set.

If $m > n$, obviously $r < m$, and it follows that any set of m linear functions in less than m unknowns must be linearly dependent.

The fact that the polynomials

$$\begin{aligned} f_1 &\equiv 2x - 3y + 4z, \\ f_2 &\equiv x + 2y - 3z, \\ f_3 &\equiv 4x + y - 2z, \end{aligned}$$

are linearly dependent can be determined by observing that the matrix of the coefficients,

$$A \equiv \begin{pmatrix} 2 & -3 & 4 \\ 1 & 2 & -3 \\ 4 & 1 & -2 \end{pmatrix},$$

is of rank 2.

35. Consistent and Inconsistent Systems of Equations. A set of equations that have at least one common solution is said to be a *consistent set* of equations. A set for which there exists no common solution is called an *inconsistent set.*

The question of consistency is frequently of practical importance. For example, in setting up problems in electrical networks, there are often more conditions than there are variables.

* DICKSON, L. E., Modern Algebraic Theories, pp. 55–60; BOCHER, M., Introduction to Higher Algebra, pp. 34–38.

This leads to a system in which there are more equations than there are unknowns. It is important to have a method for testing whether all the conditions can be satisfied simultaneously.

THEOREM 1. *Consider a system of m linear equations in n unknowns,*

$$(35\text{-}1)\qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = k_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = k_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots, \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = k_m, \end{cases}$$

where at least one $k_i \neq 0$. If the matrix of the coefficients is of rank r, Eqs. (35-1) *are consistent provided that the rank of*

$$K \equiv \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & k_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & k_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & k_m \end{pmatrix}$$

is also r.

The matrix K is called the *augmented matrix*. The proof of this theorem will be found in any standard work on higher algebra.

Example 1. Consider the system

$$\begin{aligned} 2x + 3y &= 1, \\ x - 2y &= 4, \\ 4x - y &= 9. \end{aligned}$$

Since

$$A \equiv \begin{pmatrix} 2 & 3 \\ 1 & -2 \\ 4 & -1 \end{pmatrix}$$

is of rank 2, the equations are consistent if

$$K \equiv \begin{pmatrix} 2 & 3 & 1 \\ 1 & -2 & 4 \\ 4 & -1 & 9 \end{pmatrix}$$

is also of rank 2. This condition is satisfied; for the determinant of K is zero, and there exists a second-order determinant of K that is different from zero.

Example 2. The system

$$\begin{aligned} 2x + 3y &= 1, \\ x - 2y &= 4, \\ 4x - y &= 6 \end{aligned}$$

is inconsistent, because

$$K = \begin{pmatrix} 2 & 3 & 1 \\ 1 & -2 & 4 \\ 4 & -1 & 6 \end{pmatrix}$$

is of rank 3, whereas the matrix A is of rank 2.

In the case in which there are n equations in n unknowns, the theorem on consistent equations shows that if the determinant of A is zero, so that the rank of A is $r < n$, then the rank of K must be r also, if the set of equations is to be consistent. If the rank of K is greater than r, the set of equations is inconsistent. This provides the analytic discussion that should accompany the geometric discussion given for $n = 3$ in Sec. 30.

If the set of equations is consistent and the rank of A is r, then it can be shown that $n - r$ of the unknowns can be given arbitrary values, and the values of the remaining r unknowns are determined uniquely in terms of those $n - r$ arbitrary values. These $n - r$ unknowns cannot be chosen at random, for the $m \times r$ matrix of the coefficients of the remaining r unknowns must have rank r if these unknowns are to be uniquely determined.

Example 3. Solve the system

$$\begin{aligned} x - y + 2z &= 3, \\ x + y - 2z &= 1, \\ x + 3y - 6z &= -1. \end{aligned}$$

Since A and K are both of rank 2, the equations are consistent. If either y or z is chosen arbitrarily the matrix of the coefficients of the remaining variables will have rank 2. If $z = k$, the equations to be solved are

$$\begin{aligned} x - y &= 3 - 2k, \\ x + y &= 1 + 2k, \\ x + 3y &= -1 + 6k. \end{aligned}$$

Solving the first two for x and y gives $x = 2$ and $y = 2k - 1$. These values are seen to satisfy the third equation. Therefore, the solutions $x = 2$, $y = 2k - 1$, $z = k$ satisfy the original system for all values of k.

The preceding discussion has dealt with non-homogeneous linear equations. In case the k_i are all zero, the system becomes the set of homogeneous equations

$$\text{(35-2)} \qquad \left\{ \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots&\cdots, \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned} \right.$$

Obviously, $x_1 = x_2 = \cdots = x_n = 0$ is a solution of (35-2). It may happen that there are other solutions. If $a_1, a_2, \cdots, a_n$ is a solution of (35-2), it is evident that $ka_1, ka_2, \cdots, ka_n$, where k is an arbitrary constant, will be a solution, also. The condition for solutions different from the $x_1 = x_2 = \cdots = x_n = 0$ solution will be stated without proof.

THEOREM 2. *The system* (35-2) *will have a solution different from the solution* $x_1 = x_2 = \cdots = x_n = 0$, *if the rank of the matrix of the coefficients is less than* n.

It follows that if the number of equations is less than the number of unknowns, that is, if $m < n$, there are always solutions other than the obvious zero solution. If $m = n$, there exist other solutions if the determinant of the square matrix of the coefficients is zero. As in the case of the non-homogeneous system, if the $m \times n$ matrix of the coefficients is of rank r, then $n - r$ of the unknowns can be specified arbitrarily and the remaining r unknowns will be uniquely determined, provided that the rank of the matrix of the remaining unknowns is r.

Example 4. Consider the system

$$\begin{aligned} 2x - y + 3z &= 0, \\ x + 2y - z &= 0, \\ 3x + 4y + z &= 0. \end{aligned}$$

Here

$$|A| \equiv \begin{vmatrix} 2 & -1 & 3 \\ 1 & 2 & -1 \\ 3 & 4 & 1 \end{vmatrix} = 10.$$

Therefore, $x = 0, y = 0, z = 0$ is the only solution.

Example 5. Consider

$$\begin{aligned} 3x - 2y &= 0, \\ x + 4y &= 0, \\ 2x - y &= 0, \end{aligned}$$

for which the matrix of the coefficients is of rank 2. Since the number of unknowns is 2, $x = 0, y = 0$ is the only solution.

Example 6. Consider

$$\begin{aligned} 2x - y + 3z &= 0, \\ 3x + 2y + z &= 0, \\ x + 3y - 2z &= 0, \\ 5x + y + 4z &= 0. \end{aligned}$$

Here,

$$A \equiv \begin{pmatrix} 2 & -1 & 3 \\ 3 & 2 & 1 \\ 1 & 3 & -2 \\ 5 & 1 & 4 \end{pmatrix},$$

which is of rank 2. Since the number of unknowns is 3, the system has solutions other than $x = 0$, $y = 0$, $z = 0$. Let $z = k$, and solve any two of the equations for x and y. If the first two are chosen, $x = -k$ and $y = k$. By substitution, it is easily verified that $x = -k$, $y = k$, $z = k$ satisfies all four equations for any choice of k.

Example 7. Consider

$$\begin{aligned} 2x - 4y + z &= 0, \\ 3x + y - 2z &= 0. \end{aligned}$$

For this system,

$$A = \begin{pmatrix} 2 & -4 & 1 \\ 3 & 1 & -2 \end{pmatrix},$$

which is of rank 2. Since the number of unknowns is greater than the number of equations, there exist other solutions. Let $z = k$, and solve the two equations for x and y. There results $x = \frac{1}{2}k$ and $y = \frac{1}{2}k$. Thus, $x = \frac{1}{2}k$, $y = \frac{1}{2}k$, $z = k$ is a solution for any choice of k.

Example 8. Consider

$$\begin{aligned} x - y + z &= 0, \\ 2x + 3y + z &= 0, \\ 3x + 2y + 2z &= 0. \end{aligned}$$

Here,

$$|A| = \begin{vmatrix} 1 & -1 & 1 \\ 2 & 3 & 1 \\ 3 & 2 & 2 \end{vmatrix} = 0.$$

Since the determinant of A is zero, there are solutions different from $x = 0$, $y = 0$, $z = 0$. Let $z = k$, and solve any two of the equations. If the first two are chosen, $x = -\frac{4}{5}k$, $y = \frac{1}{5}k$, $z = k$. It is verifiable by substitution that these values satisfy all three equations, whatever be the choice of k.

PROBLEMS

1. Investigate the following systems and find solutions whenever the systems are consistent:

(*a*)
$$\begin{aligned} x - 2y &= 3, \\ 2x + y &= 1, \\ 3x - y &= 4. \end{aligned}$$

(*b*)
$$\begin{aligned} 2x + y - z &= 1, \\ x - 2y + z &= 3, \\ 4x - 3y + z &= 5. \end{aligned}$$

(*c*)
$$\begin{aligned} 3x + 2y &= 4, \\ x - 3y &= 1, \\ 2x + 5y &= -1. \end{aligned}$$

(*d*)
$$\begin{aligned} 2x - y + 3z &= 4, \\ x + y - 3z &= -1, \\ 5x - y + 3z &= 7. \end{aligned}$$

2. Investigate for consistency, and obtain non-zero solutions when they exist.

(*a*) $x + 3y - 2z = 0,$
$2x - y + z = 0.$

(*b*) $x - 2y = 0,$
$3x + y = 0,$
$2x - y = 0.$

(*c*) $3x - 2y + z = 0,$
$x + 2y - 2z = 0,$
$2x - y + 2z = 0.$

(*d*) $2x - 4y + 3z = 0,$
$x + 2y - 2z = 0,$
$3x - 2y + z = 0.$

(*e*) $4x - 2y + z = 0,$
$2x - y + 3z = 0,$
$2x - y - 2z = 0,$
$6x - 3y + 4z = 0.$

(*f*) $x + 2y + 2z = 0,$
$3x - y + z = 0,$
$2x + 3y + 2z = 0,$
$x + 4y - 2z = 0.$

CHAPTER IV

PARTIAL DIFFERENTIATION

36. Functions of Several Variables. Most of the functions considered in the preceding chapters depended on a single independent variable. This chapter is devoted to a study of functions depending on more than one independent variable.

A simple example of a function of two independent variables x and y is

$$z = xy,$$

which can be thought to represent the area of a rectangle whose sides are x and y. Again, the volume v of a rectangular parallelepiped whose edges are x, y, and z, namely,

$$v = xyz,$$

is an example of a function of three independent variables x, y, and z. A function u of n independent variables $x_1, x_2, \cdots, x_n$ can be denoted by

$$u = f(x_1, x_2, \cdots, x_n).$$

A real function of a single independent variable x, say $y = f(x)$, can be represented graphically by a curve in the xy-plane. Analogously, a real function $z = f(x, y)$, of two independent variables x and y, can be thought to represent a surface in the three-dimensional space referred to a set of coordinate axes x, y, z (Fig. 27). However, one must not become too much dependent on geometric interpretations, for such interpretations may prove to be of more hindrance than help. For instance, the function $v = xyz$, representing the volume of a rectangular parallelepiped, depends on three independent variables x, y,

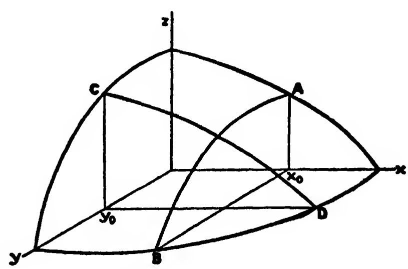

FIG. 27.

and z and hence cannot be conveniently represented geometrically in a space of three dimensions.

Corresponding to the definition of continuity of a function of a single independent variable x (see Sec. 7), it will be said that a function $z = f(x, y)$ is continuous at the point (x_0, y_0) provided that a small change in the values of x and y produces a small change in the value of z. More precisely, if the value of the function $z = f(x, y)$ at the point (x_0, y_0) is z_0, then the continuity of the function at the point (x_0, y_0) means that*

$$\lim_{\substack{x \to x_0 \\ y \to y_0}} f(x, y) = f(x_0, y_0) = z_0. \tag{36-1}$$

In writing the left-hand member of (36-1), it is assumed that the limit is independent of the mode of approach of (x, y) to (x_0, y_0).

The statement embodied in (36-1) is another way of saying that

$$f(x, y) = f(x_0, y_0) + \epsilon,$$

where $\lim_{\substack{x \to x_0 \\ y \to y_0}} \epsilon = 0$; that is, *if the function $f(x, y)$ is continuous at (x_0, y_0), then its value in the neighborhood of the point (x_0, y_0) can be made to differ from the value at the point (x_0, y_0) by as little as desired.*

If a function is continuous at all points of some region R in the xy-plane, then it is said to be *continuous in the region R.*

The definition of continuity of a function of more than two independent variables is similar. Thus, the continuity of the function $u = f(x, y, z)$ at the point (x_0, y_0, z_0) means that

$$\lim_{\substack{x \to x_0 \\ y \to y_0 \\ z \to z_0}} f(x, y, z) = f(x_0, y_0, z_0),$$

independently of the way in which (x, y, z) approaches (x_0, y_0, z_0).

PROBLEM

Describe the surfaces represented by the following equations:

(*a*) $x + 2y = 3$, (*b*) $x - y + z = 1$, (*c*) $x = 2$, (*d*) $z = y$,
(*e*) $2x - 3y + 7z = 1$, (*f*) $x^2 - y^2 = 0$, (*g*) $y^2 + z^2 = 25$,
(*h*) $y^2 = 2x$, (*i*) $x^2 + y^2 - 10x = 0$, (*j*) $x^2 + y^2 + z^2 = 1$,
(*k*) $x^2 + z^2 = y$, (*l*) $x^2 + 2y^2 + z = 0$, (*m*) $x^2 + y^2 = z^2$,

* For details, see I. S. Sokolnikoff, Advanced Calculus, Chap. III.

(n) $\dfrac{x^2}{9} + \dfrac{y^2}{4} - \dfrac{z^2}{2} = 1$, (o) $\dfrac{x^2}{9} - \dfrac{y^2}{4} - \dfrac{z^2}{2} = 1$,

(p) $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$, (q) $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 2cz$,

(r) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 2cz$, (s) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = c^2z^2$.

37. Partial Derivatives. The analytical definition of the derivative of a function $y = f(x)$, of a single variable x, is

$$\frac{df}{dx} \equiv \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

This derivative can be interpreted geometrically as the slope of the curve represented by the equation $y = f(x)$ (Fig. 28).

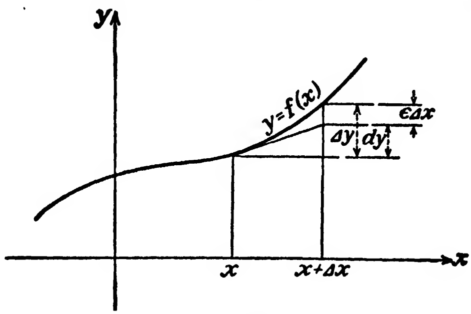

Fig. 28.

It is natural to extend the definition of the derivative to functions of several variables in the following way: Consider the function $z = f(x, y)$ of two independent variables x and y. If y is held fast, z becomes a function of the single variable x and its derivative with respect to x can be computed in the usual way. Let Δz_x denote the increment in the function $z = f(x, y)$ when y is kept fixed and x is changed by an amount Δx; that is,

$$\Delta z_x = f(x + \Delta x, y) - f(x, y).$$

Then,

$$\lim_{\Delta x \to 0} \frac{\Delta z_x}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

is called the *partial derivative of z with respect to x* and is denoted by the symbol $\partial z/\partial x$, or z_x, or f_x.

Similarly, the partial derivative of z with respect to y is defined by

$$\frac{\partial z}{\partial y} \equiv \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}.$$

In general, if $u = f(x_1, x_2, \cdots, x_n)$ is a function of n independent variables $x_1, x_2, \cdots, x_n$, then $\partial u/\partial x_i$ denotes the derivative of u with respect to x_i when the remaining variables are treated as constants. Thus, if

$$z = x^3 + x^2y + y^3,$$

then

$$\frac{\partial z}{\partial x} = 3x^2 + 2xy \qquad \text{and} \qquad \frac{\partial z}{\partial y} = x^2 + 3y^2.$$

Also, if $u = \sin (ax + by + cz)$, then

$$\frac{\partial u}{\partial x} = a \cos (ax + by + cz) \text{ (both } y \text{ and } z \text{ held constant);}$$

$$\frac{\partial u}{\partial y} = b \cos (ax + by + cz) \text{ (both } x \text{ and } z \text{ held constant);}$$

$$\frac{\partial u}{\partial z} = c \cos (ax + by + cz) \text{ (both } x \text{ and } y \text{ held constant).}$$

In the case of $z = f(x, y)$, it is easy to provide a simple geometric interpretation of partial derivatives (Fig. 27). The equation $z = f(x, y)$ is the equation of a surface; and if x is given the fixed value x_0, $z = f(x_0, y)$ is the equation of the curve AB on the surface, formed by the intersection of the surface and the plane $x = x_0$. Then $\partial z/\partial y$ gives the value of the slope at any point of AB. Similarly, if y is given the constant value y_0, then $z = f(x, y_0)$ is the equation of the curve CD on the surface, and $\partial z/\partial x$ gives the slope at any point of CD.

PROBLEMS

1. Find $\partial z/\partial x$ and $\partial z/\partial y$ for each of the following functions:

(a) $z = y/x$; (b) $z = x^3y + \tan^{-1} (y/x)$; (c) $z = \sin xy + x$;
(d) $z = e^x \log y$; (e) $z = x^2y + \sin^{-1} x$.

2. Find $\partial u/\partial x$, $\partial u/\partial y$, and $\partial u/\partial z$ for each of the following functions:

(a) $u = x^2y + yz - xz^2$; (b) $u = xyz + \log xy$;
(c) $u = z \sin^{-1} (x/y)$; (d) $u = (x^2 + y^2 + z^2)^{1/2}$;
(e) $u = (x^2 + y^2 + z^2)^{-1/2}$.

38. Total Differential. In the case of a function of one variable, $y = f(x)$, the derivative of y with respect to x is defined as

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} \equiv f'(x),$$

so that $\Delta y/\Delta x = f'(x) + \epsilon$, where $\lim\limits_{\Delta x \to 0} \epsilon = 0$. Therefore,

$$f(x + \Delta x) - f(x) \equiv \Delta y = f'(x)\,\Delta x + \epsilon\,\Delta x,$$

where ϵ is an infinitesimal which vanishes with Δx. Then,

$$f'(x)\,\Delta x \equiv f'(x)\,dx$$

is defined as the differential dy.

For the independent variable x, the terms "increment" and "differential" are synonymous (that is, $\Delta x \equiv dx$). However, it should be noted that the differential dy (of the dependent variable y) and the increment Δy differ by an amount $\epsilon\,\Delta x$ (see Fig. 28).

The differential of a function of several independent variables is defined similarly. Let $z = f(x, y)$, and let x and y acquire the respective increments Δx and Δy. Then,

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y).$$

If $z = f(x, y)$ is a continuous function, then, as Δx and Δy approach zero in any manner, Δz also approaches zero as a limit. It will be assumed here that $f(x, y)$ is continuous and that $\partial f/\partial x$ and $\partial f/\partial y$ are also continuous.

The expression for Δz can be put in a more useful form by adding and subtracting the term $f(x, y + \Delta y)$. Then,

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) + f(x, y + \Delta y) - f(x, y).$$

But

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)}{\Delta x} = \frac{\partial f(x, y + \Delta y)}{\partial x},$$

so that

$$f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) = \left[\frac{\partial f(x, y + \Delta y)}{\partial x} + \epsilon_1\right] \Delta x,$$

where $\lim\limits_{\Delta x \to 0} \epsilon_1 = 0$. Moreover,

$$\lim_{\Delta y \to 0} \frac{\partial f(x, y + \Delta y)}{\partial x} = \frac{\partial f(x, y)}{\partial x},$$

since the derivative is continuous. Therefore,

$$\frac{\partial f(x, y + \Delta y)}{\partial x} = \frac{\partial f(x, y)}{\partial x} + \epsilon_2,$$

where $\lim\limits_{\Delta y \to 0} \epsilon_2 = 0$.

In like manner,

$$f(x, y + \Delta y) - f(x, y) = \left[\frac{\partial f(x, y)}{\partial y} + \epsilon'\right] \Delta y,$$

where $\lim\limits_{\Delta y \to 0} \epsilon' = 0$. It follows that

$$\Delta z = \frac{\partial f(x, y)}{\partial x} \Delta x + \frac{\partial f(x, y)}{\partial y} \Delta y + \epsilon \, \Delta x + \epsilon' \, \Delta y,$$

in which $\epsilon = \epsilon_1 + \epsilon_2$.

The expression

$$\frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y \equiv \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

is defined as the *total differential* of z and denoted by dz. In general, if $u = f(x_1, x_2, \cdots, x_n)$, the total differential is given by

$$du = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n. \tag{38-1}$$

The expression for the total differential is called the *principal part* of the increment Δu, and is a close approximation to Δu for sufficiently small values of $dx_1, dx_2, \cdots$, and dx_n. As in the case of a function of a single independent variable, the differential of each independent variable is identical with the increment of that variable, but the differential of the dependent variable differs from the increment.

If all of the variables except one, say x_i, are considered as constants, the resulting differential is called the partial differential and is denoted by

$$d_{x_i} u = \frac{\partial f}{\partial x_i} dx_i.$$

The partial differential expresses, approximately, the change in u due to a change $\Delta x_i \equiv dx_i$ in the independent variable x_i. On the other hand, the total differential du expresses, approximately, the change in u due to changes $dx_1, dx_2, \cdots, dx_n$ in all

the independent variables $x_1, x_2, \cdots, x_n$. It may be noted that the total differential is equal to the sum of the partial differentials. Physically, this corresponds to the principle of superposition of effects. When a number of changes are taking place simultaneously in any system, each one proceeds as if it were independent of the others and the total change is the sum of the effects due to the independent changes.

Example 1. A metal box without a top has inside dimensions $6 \times 4 \times 2$ ft. If the metal is 0.1 ft. thick, find the actual volume of the metal used and compare it with the approximate volume found by using the differential.

The actual volume is ΔV, where

$$\Delta V = 6.2 \times 4.2 \times 2.1 - 6 \times 4 \times 2 = 54.684 - 48 = 6.684 \text{ cu. ft.}$$

Since $V = xyz$, where $x = 6$, $y = 4$, $z = 2$,

$$\begin{aligned} dV &= yz\,dx + xz\,dy + xy\,dz \\ &= 8(0.2) + 12(0.2) + 24(0.1) = 6.4 \text{ cu. ft.} \end{aligned}$$

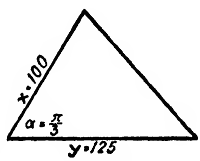

FIG. 29.

Example 2. Two sides of a triangular piece of land (Fig. 29) are measured as 100 ft. and 125 ft., and the included angle is measured as 60°. If the possible errors are 0.2 ft. in measuring the sides and 1° in measuring the angle, what is the approximate error in the area?

Since $A = \frac{1}{2}xy \sin \alpha$,

$$dA = \tfrac{1}{2}(y \sin \alpha\, dx + x \sin \alpha\, dy + xy \cos \alpha\, d\alpha),$$

and the approximate error is therefore

$$dA = \frac{1}{2}\left[125\left(\frac{\sqrt{3}}{2}\right)(0.2) + 100\left(\frac{\sqrt{3}}{2}\right)(0.2) + 100(125)\left(\frac{1}{2}\right)\frac{\pi}{180}\right] = 74.0 \text{ sq. ft.}$$

PROBLEMS

1. A closed cylindrical tank is 4 ft. high and 2 ft. in diameter (inside dimensions). What is the approximate amount of metal in the wall and the ends of the tank if they are 0.2 in. thick?

2. The angle of elevation of the top of a tower is found to be 30°, with a possible error of 0.5°. The distance to the base of the tower is found to be 1000 ft., with a possible error of 0.1 ft. What is the possible error in the height of the tower as computed from these measurements?

3. What is the possible error in the length of the hypotenuse of a right triangle if the legs are found to be 11.5 ft. and 7.8 ft., with a possible error of 0.1 ft. in each measurement?

4. The constant C in Boyle's law $pv = C$ is calculated from the measurements of p and v. If p is found to be 5000 lb. per square foot with a possible error of 1 per cent and v is found to be 15 cu. ft. with a possible error of 2 per cent, find the approximate possible error in C computed from these measurements.

5. The volume v, pressure p, and absolute temperature T of a perfect gas are connected by the formula $pv = RT$, where R is a constant. If $T = 500°$, $p = 4000$ lb. per square foot, and $v = 15.2$ cu. ft., find the approximate change in p when T changes to 503° and v to 15.25 cu. ft.

6. In estimating the cost of a pile of bricks measured as $6 \times 50 \times 4$ ft., the tape is stretched 1 per cent beyond the estimated length. If the count is 12 bricks to 1 cu. ft. and bricks cost \$8 per thousand, find the error in cost.

7. In determining specific gravity by the formula $s = \dfrac{A}{A - W}$, where A is the weight in air and W is the weight in water, A can be read within 0.01 lb. and W within 0.02 lb. Find approximately the maximum error in s if the readings are $A = 1.1$ lb. and $W = 0.6$ lb. Find the maximum relative error $\Delta s/s$.

8. The equation of a perfect gas is $pv = RT$. At a certain instant a given amount of gas has a volume of 16 cu. ft. and is under a pressure of 36 lb. per square inch. Assuming $R = 10.71$, find the temperature T. If the volume is increasing at the rate of $\frac{1}{3}$ cu. ft. per second and the pressure is decreasing at the rate $\frac{1}{6}$ lb. per square inch per second, find the rate at which the temperature is changing.

9. The period of a simple pendulum with small oscillations is

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

If T is computed using $l = 8$ ft. and $g = 32$ ft. per second per second, find the approximate error in T if the true values are $l = 8.05$ ft. and $g = 32.01$ ft. per second per second. Find also the percentage error.

10. The diameter and altitude of a can in the shape of a right circular cylinder are measured as 4 in. and 6 in., respectively. The possible error in each measurement is 0.1 in. Find approximately the maximum possible error in the values computed for the volume and the lateral surface.

39. Total Derivatives. Thus far, it has been assumed that x and y were independent variables. It may be that x and y

are both functions of one independent variable t, so that z becomes a function of this single independent variable. In such a case, z may have a derivative with respect to t.

Let $z = f(x, y)$, where $x = \varphi(t)$ and $y = \psi(t)$; these functions are assumed to be differentiable. If t is given an increment Δt, then x, y, and z will have corresponding increments Δx, Δy, and Δz, which approach zero with Δt. As in the case when x and y were independent variables,

$$\Delta z = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y.$$

Then,

$$\frac{\Delta z}{\Delta t} = \frac{\partial f}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y}\frac{\Delta y}{\Delta t} + \epsilon_1 \frac{\Delta x}{\Delta t} + \epsilon_2 \frac{\Delta y}{\Delta t}$$

and

(39-1) $$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Moreover, from (39-1) it appears that

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

gives the expression for the differential in this case as well as when x and y are independent variables.

The general case, in which

$$z = f(x_1, x_2, \cdots, x_n)$$

with

$$x_1 = \varphi_1(t), \qquad x_2 = \varphi_2(t), \cdots, \qquad x_n = \varphi_n(t),$$

can be treated similarly to show that

$$\frac{dz}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}$$

and

$$dz = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n.$$

In case $t = x$ (39-1) becomes

$$\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} \equiv \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx}.$$

This formula can be used to calculate the derivative of a func-

tion of x defined implicitly by the equation $f(x, y) = 0$. Let $z = f(x, y)$, so that

$$\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}.$$

Since $z = f(x, y) = 0$, it follows that $dz/dx = 0$ and

$$\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y},$$

provided that $\partial f/\partial y \neq 0$.

As an example, let

$$x^2y + y^2x - 1 = 0$$

define y as an implicit function of x. Then,

$$\frac{dy}{dx} = -\frac{2xy + y^2}{x^2 + 2xy}$$

for all values of x and y for which the denominator does not vanish.

It was noted that the total differential of a function

$$z = f(x_1, x_2, \cdots, x_n), \quad \text{where } x_i = \varphi_i(t),$$

is given by

$$dz = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n.$$

It will be proved next that the same formula can be applied to calculate the differential even when the variables x_i are functions of several independent variables $t_1, t_2, \cdots, t_m$. Thus, consider

$$z = f(x_1, x_2, \cdots, x_n), \qquad \text{where } x_i = \varphi_i(t_1, t_2, \cdots, t_m).$$

In order to find the partial derivative of $f(x_1, x_2, \cdots, x_n)$ with respect to one of the variables, say t_k, the remaining variables are held fixed so that $f(x_1, x_2, \cdots, x_n)$ becomes a function of the single variable t_k. Then,

$$(39\text{-}2)\qquad \begin{cases} \dfrac{\partial f}{\partial t_1} = \dfrac{\partial f}{\partial x_1}\dfrac{\partial x_1}{\partial t_1} + \dfrac{\partial f}{\partial x_2}\dfrac{\partial x_2}{\partial t_1} + \cdots + \dfrac{\partial f}{\partial x_n}\dfrac{\partial x_n}{\partial t_1}, \\ \dfrac{\partial f}{\partial t_2} = \dfrac{\partial f}{\partial x_1}\dfrac{\partial x_1}{\partial t_2} + \dfrac{\partial f}{\partial x_2}\dfrac{\partial x_2}{\partial t_2} + \cdots + \dfrac{\partial f}{\partial x_n}\dfrac{\partial x_n}{\partial t_2}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{\partial f}{\partial t_m} = \dfrac{\partial f}{\partial x_1}\dfrac{\partial x_1}{\partial t_m} + \dfrac{\partial f}{\partial x_2}\dfrac{\partial x_2}{\partial t_m} + \cdots + \dfrac{\partial f}{\partial x_n}\dfrac{\partial x_n}{\partial t_m}. \end{cases}$$

If the first equation of (39-2) is multiplied by dt_1, the second by dt_2, etc., and the resulting equations are added, there results

$$
\begin{aligned}
\frac{\partial f}{\partial t_1}\,dt_1 + \frac{\partial f}{\partial t_2}\,dt_2 + \cdots + \frac{\partial f}{\partial t_m}\,dt_m
&= \frac{\partial f}{\partial x_1}\left(\frac{\partial x_1}{\partial t_1}\,dt_1 + \frac{\partial x_1}{\partial t_2}\,dt_2 + \cdots + \frac{\partial x_1}{\partial t_m}\,dt_m\right)\\
&+ \frac{\partial f}{\partial x_2}\left(\frac{\partial x_2}{\partial t_1}\,dt_1 + \frac{\partial x_2}{\partial t_2}\,dt_2 + \cdots + \frac{\partial x_2}{\partial t_m}\,dt_m\right)\\
&+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&+ \frac{\partial f}{\partial x_n}\left(\frac{\partial x_n}{\partial t_1}\,dt_1 + \frac{\partial x_n}{\partial t_2}\,dt_2 + \cdots + \frac{\partial x_n}{\partial t_m}\,dt_m\right)
\end{aligned}
$$

or

$$
df = \frac{\partial f}{\partial x_1}\,dx_1 + \frac{\partial f}{\partial x_2}\,dx_2 + \cdots + \frac{\partial f}{\partial x_n}\,dx_n. \tag{39-3}
$$

This establishes the validity of the formula (38-1) in all cases where the first partial derivatives are continuous functions, irrespective of whether the independent variables are $x_1, x_2, \cdots, x_n$ or $t_1, t_2, \cdots, t_m$.

An important special case of the formula (39-3) arises in aerodynamics and other branches of applied mathematics. Consider a function $u = f(x, y, z, t)$ of four variables x, y, z, and t. The total differential of u is

$$
\begin{aligned}
du &= \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz + \frac{\partial f}{\partial t}\,dt\\
&\equiv \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy + \frac{\partial u}{\partial z}\,dz + \frac{\partial u}{\partial t}\,dt.
\end{aligned} \tag{39-4}
$$

Let it be supposed that x, y, and z are not independent variables, but functions of the variable t. In such a case, u will depend on t explicitly, and also implicitly through x, y, and z. Dividing both members of (39-4) by dt gives

$$
\frac{du}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt} + \frac{\partial u}{\partial z}\frac{dz}{dt} + \frac{\partial u}{\partial t}. \tag{39-5}
$$

On the other hand, if the variables x, y, and z are functions of t and of some other set of independent variables $r, s, \cdots$, one must replace dx/dt, dy/dt, and dz/dt in the right-hand member of (39-5) by $\partial x/\partial t$, $\partial y/\partial t$, and $\partial z/\partial t$, respectively, and du/dt in

the left-hand member by $\partial u/\partial t$. The partial derivative with respect to t which appears in the left-hand member differs from that appearing in the right-hand member, since the latter is computed from $u = f(x, y, z, t)$ by fixing the variables x, y, and z and differentiating the resulting function with respect to t. In order to indicate the distinction between the meanings of the two partial derivatives with respect to t, one can write

$$\frac{Du}{\partial t} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial t} + \frac{\partial u}{\partial t}.$$

The fact that the total differential of a composite function has the same form irrespective of whether the variables involved are independent or not permits one to use the same formulas for calculating differentials as those established for the functions of a single variable. Thus,

$$\begin{aligned} d(u + v) &= du + dv, \\ d(uv) &= \frac{\partial(uv)}{\partial u}\,du + \frac{\partial(uv)}{\partial v}\,dv \\ &= v\,du + u\,dv, \end{aligned}$$

etc.

Example 1. If $u = xy + yz + zx$, and $x = t$, $y = e^{-t}$, and $z = \cos t$,

$$\begin{aligned} \frac{du}{dt} &= (y + z)\frac{dx}{dt} + (x + z)\frac{dy}{dt} + (x + y)\frac{dz}{dt} \\ &= (e^{-t} + \cos t)(1) + (t + \cos t)(-e^{-t}) + (t + e^{-t})(-\sin t) \\ &= e^{-t} + \cos t - te^{-t} - e^{-t}\cos t - t\sin t - e^{-t}\sin t: \end{aligned}$$

This example illustrates the fact that this method of computing du/dt is often shorter than the old method in which the values of x, y, and z in terms of t are substituted in the expression for u before the derivative is computed.

Example 2. If $f(x, y) = x^2 + y^2$, where $x = r\cos\varphi$ and $y = r\sin\varphi$, then

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = 2x\cos\varphi + 2y\sin\varphi = 2r\cos^2\varphi + 2r\sin^2\varphi = 2r.$$

$$\begin{aligned} \frac{\partial f}{\partial \varphi} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \varphi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \varphi} &= 2x(-r\sin\varphi) + 2y(r\cos\varphi) \\ &= -2r^2\cos\varphi\sin\varphi + 2r^2\cos\varphi\sin\varphi = 0. \end{aligned}$$

Also,

$$df = 2r\,dr \qquad \text{or} \qquad df = 2x\,dx + 2y\,dy.$$

Example 3. Let $z = e^{xy}$, where $x = \log(u + v)$ and $y = \tan^{-1}(u/v)$. Then,

$$\frac{\partial z}{\partial x} = ye^{xy}, \qquad \frac{\partial z}{\partial y} = xe^{xy}, \qquad \frac{\partial x}{\partial u} = \frac{1}{u+v}, \qquad \text{and} \qquad \frac{\partial y}{\partial u} = \frac{v}{v^2+u^2}.$$

Hence,

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} = \frac{ye^{xy}}{u+v} + \frac{xe^{xy}v}{v^2+u^2}.$$

Similarly,

$$\frac{\partial z}{\partial v} = \frac{ye^{xy}}{u+v} - \frac{xe^{xy}u}{v^2+u^2}.$$

The same results can be obtained by noting that

$$dz = ye^{xy}\,dx + xe^{xy}\,dy.$$

But

$$dx = \frac{\partial x}{\partial u}\,du + \frac{\partial x}{\partial v}\,dv = \frac{1}{u+v}\,du + \frac{1}{u+v}\,dv$$

and

$$dy = \frac{\partial y}{\partial u}\,du + \frac{\partial y}{\partial v}\,dv = \frac{v}{v^2+u^2}\,du - \frac{u}{v^2+u^2}\,dv.$$

Hence,

$$\begin{aligned} dz &= ye^{xy}\frac{du+dv}{u+v} + xe^{xy}\frac{v\,du - u\,dv}{v^2+u^2} \\ &= \left(\frac{ye^{xy}}{u+v} + \frac{xe^{xy}v}{v^2+u^2}\right)du + \left(\frac{ye^{xy}}{u+v} - \frac{xe^{xy}u}{v^2+u^2}\right)dv. \end{aligned}$$

But

$$dz = \frac{\partial z}{\partial u}\,du + \frac{\partial z}{\partial v}\,dv;$$

and since du and dv are independent differentials, equating the coefficients of du and dv in the two expressions for dz gives

$$\frac{\partial z}{\partial u} = \frac{ye^{xy}}{u+v} + \frac{xe^{xy}v}{v^2+u^2}$$

and

$$\frac{\partial z}{\partial v} = \frac{ye^{xy}}{u+v} - \frac{xe^{xy}u}{v^2+u^2}.$$

PROBLEMS

1. If $u = xyz$ and $x = a\cos\theta$, $y = a\sin\theta$, $z = k\theta$, find $du/d\theta$.

2. If $u = x^2 - y^2$ and $y = r\sin\theta$ and $x = r\cos\theta$, find $\partial u/\partial r$ and $\partial u/\partial\theta$.

3. If $u = xy - yz$ and $x = r + s$, $y = r - s$, $z = t$, find $\partial u/\partial r$ $\partial u/\partial s$, and $\partial u/\partial t$.

4. If $z = e^{xy}$, $x = \log \sqrt{u^2 + v^2}$, and $y = \tan^{-1} \dfrac{u}{v}$, find $\partial z/\partial u$ and $\partial z/\partial v$.

5. If $z = f(x + u, y + v)$, show that $\partial z/\partial x = \partial z/\partial u$ and $\partial z/\partial y = \partial z/\partial v$.

6. If $u = x^2y + y^2z + z^2x$, verify that

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} + \frac{\partial u}{\partial z} = (x + y + z)^2.$$

7. (*a*) Find du/dt, if $u = e^x \sin yz$ and $x = t^2$, $y = t - 1$, $z = 1/t$. (*b*) Find $\partial u/\partial r$ and $\partial u/\partial \theta$, if $u = x^2 - 4y^2$ and $x = r \sec \theta$, $y = r \tan \theta$.

8. (*a*) Find $\partial u/\partial x$ and du/dx, if $u = x^2 + y^2$ and $y = \tan x$. (*b*) Given $V = f(x, y, z)$, where $x = r \cos \theta$, $y = r \sin \theta$, $z = t$. Compute $\partial V/\partial r$, $\partial V/\partial \theta$, $\partial V/\partial t$ in terms of $\partial V/\partial x$, $\partial V/\partial y$, and $\partial V/\partial z$.

9. If f is a function of u and v, where $u = \sqrt{x^2 + y^2}$ and $v = \tan^{-1} \dfrac{y}{x}$, find $\partial f/\partial x$, $\partial f/\partial y$, and $\sqrt{(\partial f/\partial x)^2 + (\partial f/\partial y)^2}$.

40. Euler's Formula. A function $f(x_1, x_2, \cdots, x_n)$ of n variables $x_1, x_2, \cdots, x_n$ is said to be *homogeneous of degree m* if the function is multiplied by λ^m when the arguments $x_1, x_2, \cdots, x_n$ are replaced by $\lambda x_1, \lambda x_2, \cdots, \lambda x_n$, respectively. For example, $f(x, y) = x^2/\sqrt{x^2 + y^2}$ is homogeneous of degree 1, because the substitution of λx for x and λy for y yields $\lambda x^2/\sqrt{x^2 + y^2}$. Again, $f(x, y) = \dfrac{1}{y} + \dfrac{\log x - \log y}{x}$ is homogeneous of degree -1, whereas $f(x, y, z) = z^2/\sqrt[3]{x^2 + y^2}$ is homogeneous of degree $\frac{4}{3}$.

There is an important theorem, due to Euler, concerning homogeneous functions.

EULER'S THEOREM. *If $u = f(x_1, x_2, \cdots, x_n)$ is homogeneous of degree m and has continuous first partial derivatives, then*

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \cdots + x_n \frac{\partial f}{\partial x_n} = mf(x_1, x_2, \cdots, x_n).$$

The proof of the theorem follows at once upon substituting

$$x'_1 = \lambda x_1, \qquad x'_2 = \lambda x_2, \quad \cdots, \quad x'_n = \lambda x_n.$$

Then, since $f(x_1, x_2, \cdots, x_n)$ is homogeneous of degree m,

$$f(x'_1, x'_2, \cdots, x'_n) = \lambda^m f(x_1, x_2, \cdots, x_n).$$

Differentiating with respect to λ gives

$$\frac{\partial f}{\partial x_1'} x_1 + \frac{\partial f}{\partial x_2'} x_2 + \cdots + \frac{\partial f}{\partial x_n'} x_n = m\lambda^{m-1} f(x_1, x_2, \cdots, x_n).$$

If λ is set equal to 1, then $x_1 = x_1'$, $x_2 = x_2'$, $\cdots$, $x_n = x_n'$ and the theorem follows.

PROBLEM

Verify Euler's theorem for each of the following functions:

(a) $f(x, y, z) = x^2y + xy^2 + 2xyz$;

(b) $f(x, y) = \sqrt{y^2 - x^2} \sin^{-1} \dfrac{x}{y}$;

(c) $f(x, y) = \dfrac{1}{y^2} + \dfrac{\log x - \log y}{x^2}$;

(d) $f(x, y, z) = \dfrac{z^2}{\sqrt{x^2 - y^2}}$;

(e) $f(x, y, z) = (x^2 + y^2 + z^2)^{-1/2}$;

(f) $f(x, y) = e^{x/y}$;

(g) $f(x, y) = \dfrac{\sqrt{x + y}}{y}$;

(h) $f(x, y) = \dfrac{x^2 + y^2}{x^2 - y^2}$.

41. Differentiation of Implicit Functions. It was noted in Sec. 39 that the derivative of a function of x which is defined implicitly by the equation $f(x, y) = 0$ could be calculated by applying the expression for the total derivative. This section contains a more detailed discussion of this method.

The equation $f(x, y) = 0$ may define either x or y as an implicit function of the other. If the equation can be solved for y to give $y = \varphi(x)$, then the substitution of $y = \varphi(x)$ in $f(x, y) = 0$ gives an identity. Hence, $f(x, y) = 0$ may be regarded as a composite function of x, where x enters implicitly in y. If $u = f(x, y)$, then

$$du = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0,$$

so that

$$\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y}, \qquad \text{if } \frac{\partial f}{\partial y} \neq 0. \tag{41-1}$$

It will be observed that this discussion tacitly assumes that $f(x, y) = 0$ has a real solution for y for every value of x. If (41-1) is applied formally to $x^2 + y^2 = 0$, it is readily checked

that $\dfrac{dy}{dx} = -\dfrac{x}{y}$. This result is absurd for real values of x and y, inasmuch as the only real values of x and y that satisfy $x^2 + y^2 = 0$ are $x = 0$ and $y = 0$.

Example 1. Find dy/dx, if $3x^3y^2 + x \cos y = 0$. Here,

$$\frac{\partial f}{\partial x} = 9x^2y^2 + \cos y, \qquad \frac{\partial f}{\partial y} = 6x^3y - x \sin y,$$

so that

$$\frac{dy}{dx} = -\frac{9x^2y^2 + \cos y}{6x^3y - x \sin y}.$$

The relation $f(x, y, z) = 0$ may define any one of the variables as an implicit function of the other two. Let x and y be independent variables. Then $f(x, y, z) = 0$ defines z as an implicit function of x and y, and

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy.$$

But

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz = 0.$$

Therefore, by substitution,

$$\frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z}\left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy\right) = 0.$$

This can be written as

$$\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y}\right) dy = 0.$$

Since dx and dy are independent differentials and the above relation holds for all values of dx and dy, it follows that

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 0$$

and

$$\frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} = 0.$$

If $\partial f/\partial z \neq 0$, these equations can be solved to give

$$\text{(41-2)} \qquad \frac{\partial z}{\partial x} = -\frac{\partial f/\partial x}{\partial f/\partial z}, \qquad \frac{\partial z}{\partial y} = -\frac{\partial f/\partial y}{\partial f/\partial z}.$$

Example 2. If $x^2 + 2y^2 - 3xz = 0$, then, by (41-2),

$$\frac{\partial z}{\partial x} = -\frac{2x - 3z}{-3x}, \qquad \frac{\partial z}{\partial y} = -\frac{4y}{-3x}.$$

Frequently, it is necessary to calculate the derivatives of a function that is defined implicitly by a pair of simultaneous equations

$$\begin{cases} f(x, y, z) = 0, \\ \varphi(x, y, z) = 0. \end{cases} \tag{41-3}$$

If each of these equations is solved for one of the variables, say z, to yield

$$z = F(x, y) \qquad \text{and} \qquad z = \Phi(x, y),$$

then one is led to consider the equation resulting from the elimination of z, namely,

$$F(x, y) - \Phi(x, y) = 0.$$

This equation may be thought to define y as an implicit function of x, and one can apply the method discussed earlier in this section to calculate dy/dx.

However, the elimination of one of the variables from the simultaneous equations (41-3) may prove to be difficult, and it is simpler to use the following procedure: The differentiation of (41-3) gives

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz = 0$$

and

$$d\varphi = \frac{\partial \varphi}{\partial x}\,dx + \frac{\partial \varphi}{\partial y}\,dy + \frac{\partial \varphi}{\partial z}\,dz = 0.$$

These equations can be solved for the ratios to give

$$dx:dy:dz = \begin{vmatrix} \dfrac{\partial f}{\partial y} & \dfrac{\partial f}{\partial z} \\ \dfrac{\partial \varphi}{\partial y} & \dfrac{\partial \varphi}{\partial z} \end{vmatrix} : \begin{vmatrix} \dfrac{\partial f}{\partial z} & \dfrac{\partial f}{\partial x} \\ \dfrac{\partial \varphi}{\partial z} & \dfrac{\partial \varphi}{\partial x} \end{vmatrix} : \begin{vmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial \varphi}{\partial x} & \dfrac{\partial \varphi}{\partial y} \end{vmatrix},$$

from which the derivatives can be written down at once.

Example 3. Let

$$f(x, y, z) \equiv x^2 + y^2 + z^2 - a^2 = 0$$

and

$$\varphi(x, y, z) \equiv x^2 - y^2 - 2z^2 - b^2 = 0.$$

Then,

$$\begin{aligned} dx:dy:dz &= \begin{vmatrix} 2y & 2z \\ -2y & -4z \end{vmatrix} : \begin{vmatrix} 2z & 2x \\ -4z & 2x \end{vmatrix} : \begin{vmatrix} 2x & 2y \\ 2x & -2y \end{vmatrix} \\ &= -4yz : 12xz : -8xy. \end{aligned}$$

Hence,

$$\frac{dy}{dx} = \frac{12xz}{-4yz}, \qquad \frac{dz}{dx} = \frac{-8xy}{-4yz}, \text{ etc.}$$

Another important case arises from a consideration of a pair of simultaneous equations

$$\text{(41-4)} \qquad \begin{cases} f(x, y, u, v) = 0, \\ \varphi(x, y, u, v) = 0, \end{cases}$$

which may be thought to define u and v as implicit functions of the variables x and y.

Differentiating (41-4) gives

$$\text{(41-5)} \qquad \begin{cases} df = \dfrac{\partial f}{\partial x}\, dx + \dfrac{\partial f}{\partial y}\, dy + \dfrac{\partial f}{\partial u}\, du + \dfrac{\partial f}{\partial v}\, dv = 0, \\[2ex] d\varphi = \dfrac{\partial \varphi}{\partial x}\, dx + \dfrac{\partial \varphi}{\partial y}\, dy + \dfrac{\partial \varphi}{\partial u}\, du + \dfrac{\partial \varphi}{\partial v}\, dv = 0. \end{cases}$$

But, since u and v are regarded as functions of x and y,

$$du = \frac{\partial u}{\partial x}\, dx + \frac{\partial u}{\partial y}\, dy$$

and

$$dv = \frac{\partial v}{\partial x}\, dx + \frac{\partial v}{\partial y}\, dy.$$

Substituting for du and dv in (41-5) gives

$$\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y}\right) dy = 0,$$

$$\left(\frac{\partial \varphi}{\partial x} + \frac{\partial \varphi}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial \varphi}{\partial v}\frac{\partial v}{\partial x}\right) dx + \left(\frac{\partial \varphi}{\partial y} + \frac{\partial \varphi}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial \varphi}{\partial v}\frac{\partial v}{\partial y}\right) dy = 0.$$

Since the variables x and y are independent, the coefficients of dx and dy must vanish, and this leads to a set of four equations for the determination of $\partial u/\partial x$, $\partial u/\partial y$, $\partial v/\partial x$, and $\partial v/\partial y$. Thus, one obtains

$$\frac{\partial u}{\partial x} = -\frac{\begin{vmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial v} \\[2ex] \dfrac{\partial \varphi}{\partial x} & \dfrac{\partial \varphi}{\partial v} \end{vmatrix}}{\begin{vmatrix} \dfrac{\partial f}{\partial u} & \dfrac{\partial f}{\partial v} \\[2ex] \dfrac{\partial \varphi}{\partial u} & \dfrac{\partial \varphi}{\partial v} \end{vmatrix}},$$

and similar expressions for $\partial u/\partial y$, $\partial v/\partial x$, and $\partial v/\partial y$. It is assumed in the foregoing discussion that all the derivatives involved are continuous and that

$$J \equiv \begin{vmatrix} \dfrac{\partial f}{\partial u} & \dfrac{\partial f}{\partial v} \\ \dfrac{\partial \varphi}{\partial u} & \dfrac{\partial \varphi}{\partial v} \end{vmatrix} \neq 0.$$

Example 4. If

$$\begin{aligned} x + y^3 + u^3 + v^3 &= 0, \\ x^3 + y - u^4 + v^4 &= 0, \end{aligned}$$

then

$$\frac{\partial u}{\partial x} = -\frac{\begin{vmatrix} 1 & 3v^2 \\ 3x^2 & 4v^3 \end{vmatrix}}{\begin{vmatrix} 3u^2 & 3v^2 \\ -4u^3 & 4v^3 \end{vmatrix}} = \frac{9x^2v^2 - 4v^3}{12(u^2v^3 + v^2u^3)}.$$

PROBLEMS

1. Obtain $\partial v/\partial x$, $\partial u/\partial y$, and $\partial v/\partial y$ in Example 4, Sec. 41.
2. Compute dy/dx, if $x^3 + y^3 - 3xy = 1$.
3. Find dy/dx if

$$\begin{aligned} x^2 - y^2 - z^2 - a^2 &= 0, \\ x^2y - y^2z + xz^2 - a^3 &= 0. \end{aligned}$$

4. Obtain $\partial u/\partial x$ and $\partial v/\partial y$, if

$$\begin{aligned} ue^v - xy + v &= 0, \\ ve^u - xv + u &= 0. \end{aligned}$$

5. If $x = f(u, v)$ and $y = g(u, v)$, then differentiation with respect to x gives

$$\begin{aligned} 1 &= \frac{\partial x}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial x}{\partial v}\frac{\partial v}{\partial x}, \\ 0 &= \frac{\partial y}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial y}{\partial v}\frac{\partial v}{\partial x}, \end{aligned}$$

from which $\partial u/\partial x$ and $\partial v/\partial x$ can be computed. Consider the pair of equations

$$\begin{aligned} x &= u^2 - v^2, \\ y &= uv, \end{aligned}$$

and obtain $\partial u/\partial x$, $\partial u/\partial y$, $\partial v/\partial x$, and $\partial v/\partial y$.

6. Apply the method outlined in Prob. 5 to find $\partial u/\partial x$, $\partial v/\partial x$, $\partial u/\partial y$, and $\partial v/\partial y$, if

$$(a)\ \begin{cases} x = u + v, \\ y = 3u + 2v; \end{cases}$$

$$(b)\ \begin{cases} 2x = v^2 - u^2, \\ y = uv. \end{cases}$$

7. If $x = r\cos\theta$ and $y = r\sin\theta$, find $\partial r/\partial x$ and $\partial\theta/\partial x$.

8. If $w = uv$ and

$$(a)\quad \begin{cases} u^2 + v + x = 0, \\ v^2 - u - y = 0, \end{cases}$$

one can obtain $\partial w/\partial x$ as follows: Differentiation of w with respect to x gives $\dfrac{\partial w}{\partial x} = u\dfrac{\partial v}{\partial x} + v\dfrac{\partial u}{\partial x}\cdot$ The values of $\partial u/\partial x$ and $\partial v/\partial x$ can be calculated from (*a*) by the method of Prob. 5. Find the expressions for $\partial w/\partial x$ and $\partial w/\partial y$.

9. If $z = uv$ and

$$u^2 + v^2 - x - y = 0,$$
$$u^2 - v^2 + 3x + y = 0,$$

find $\partial z/\partial x$.

10. If $z = u^2 + v^2$ and

$$x = u^2 - v^2,$$
$$y = uv,$$

find $\partial z/\partial x$.

11. If $z = u^2 + v^2$ and

$$u = r\cos\theta,$$
$$v = r\sin\theta,$$

find $\partial z/\partial r$ and $\partial z/\partial\theta$.

12. If $r = (x^2 + y^2)^{1/2}$ and $\theta = \tan^{-1}\dfrac{y}{x}$, find $\partial r/\partial x$ and $\partial\theta/\partial x$.

13. (*a*) Find dy/dx, if $x\sec y + x^3y^2 = 0$.

(*b*) Find $\partial z/\partial x$ and $\partial z/\partial y$, if $x^3y - \sin z + z^3 = 0$.

14. Let $u \equiv x + y + z = 0$ and $v \equiv x^2 + y^2 + z^2 - a^2 = 0$. Find $dx{:}dy{:}dz$.

15. Find $\partial u/\partial x$, $\partial v/\partial x$, $\partial u/\partial y$, and $\partial v/\partial y$, if

$$u^2 + v^2 + y^2 - 2x = 0,$$
$$u^3 + v^3 - x^3 + 3y = 0.$$

16. Find $\partial w/\partial x$ and $\partial w/\partial y$, if $w = u/v$ and

$$x = u + v,$$
$$y = 3u + 2v.$$

17. Show that $\dfrac{\partial z}{\partial x}\dfrac{\partial x}{\partial z} = 1$ and $\dfrac{\partial x}{\partial y}\dfrac{\partial y}{\partial z}\dfrac{\partial z}{\partial x} = -1$, if $f(x, y, z) = 0$.
Note that, in general, $\partial z/\partial x$ and $\partial x/\partial z$ are not reciprocals.

18. Find $\partial u/\partial x$, if

$$u^2 - v^2 - x^3 + 3y = 0,$$
$$u + v - y^2 - 2x = 0.$$

19. Prove that

$$\frac{\partial u}{\partial x}\frac{\partial y}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial y}{\partial v} = 0,$$

if $F(x, y, u, v) = 0$ and $G(x, y, u, v) = 0$.

20. Show that $\left(\frac{\partial z}{\partial r}\right)^2 + 1/r^2 \left(\frac{\partial z}{\partial \theta}\right)^2 = \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2$, if $x = r \cos \theta$ and $y = r \sin \theta$.

42. Directional Derivatives. The relation expressed in (39-1) has an important special case when x and y are functions of the distance s along some curve C, which goes through the point (x, y). The curve C may be thought to be represented by a pair of parametric equations

$$x = x(s),$$
$$y = y(s),$$

where x and y are assumed to possess continuous derivatives with respect to the arc parameter s.

Let P (Fig. 30) be any point of the curve C at which $f(x, y)$ is defined and has partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$. Let

$$Q(x + \Delta x, y + \Delta y)$$

be a point close to P on this curve. If Δs is the length of the arc PQ and Δf is the change in f due to the increments Δx and Δy, then

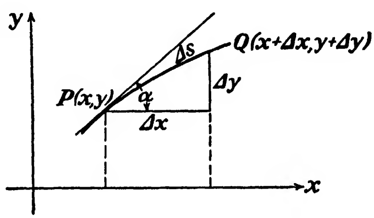

FIG. 30.

$$\frac{df}{ds} = \lim_{\Delta s \to 0} \frac{\Delta f}{\Delta s}$$

gives the rate of change of f along C at the point (x, y). But

$$\frac{df}{ds} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds},$$

and

$$\frac{dx}{ds} = \lim_{\Delta s \to 0} \frac{\Delta x}{\Delta s} = \cos \alpha, \qquad \frac{dy}{ds} = \lim_{\Delta s \to 0} \frac{\Delta y}{\Delta s} = \sin \alpha.$$

Therefore,

$$\frac{df}{ds} = \frac{\partial f}{\partial x} \cos \alpha + \frac{\partial f}{\partial y} \sin \alpha, \tag{42-1}$$

and it is evident that df/ds depends on the direction of the curve. For this reason, df/ds is called the *directional derivative*. It represents the rate of change of f in the direction of the tangent to

the particular curve chosen for the point (x, y). If $\alpha = 0$,

$$\frac{df}{ds} = \frac{\partial f}{\partial x},$$

which is the rate of change of f in the direction of the x-axis. If $\alpha = \pi/2$,

$$\frac{df}{ds} = \frac{\partial f}{\partial y},$$

which is the rate of change of f in the direction of the y-axis.

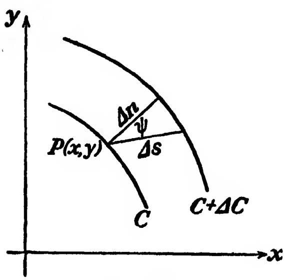

Fig. 31.

Let $z = f(x, y)$, which can be interpreted as the equation of a surface, be represented by drawing the contour lines on the xy-plane for various values of z. Let C (Fig. 31) be the curve in the xy-plane corresponding to the value $z = \gamma$, and let $C + \Delta C$ be the neighboring contour line for $z = \gamma + \Delta\gamma$. Then, $\Delta f/\Delta s \equiv \Delta\gamma/\Delta s$ is the average rate of change of f with respect to the distance Δs between C and $C + \Delta C$. Apart from infinitesimals of higher order,

$$\frac{\Delta n}{\Delta s} = \cos \psi,$$

where Δn denotes the distance from C to $C + \Delta C$ along the normal to C at (x, y), and ψ is the angle between Δn and Δs; hence, $dn/ds = \cos \psi$. Therefore,

$$\frac{df}{ds} = \frac{df}{dn} \cdot \frac{dn}{ds} = \frac{df}{dn} \cos \psi. \tag{42-2}$$

This relation shows that the derivative of f in any direction may be found by multiplying the derivative along the normal by the cosine of the angle ψ between the particular direction and the normal. This derivative in the direction of the normal is called the *normal derivative* of f. Its numerical value obviously is the maximum value that df/ds can take for any direction. In applied mathematics the vector in the direction of the normal, of magnitude df/dn, is called the *gradient*.

Example. Using (42-1), find the value of α that makes df/ds a maximum, considering x and y to be fixed. Find the expression for this maximum value of df/ds.

Since $df/ds = f_x \cos\alpha + f_y \sin\alpha$,

$$\frac{d}{d\alpha}\left(\frac{df}{ds}\right) = -f_x \sin\alpha + f_y \cos\alpha.$$

The condition for a maximum requires that

$$\tan\alpha_1 = \frac{f_y}{f_x}, \qquad \text{or} \qquad \alpha_1 = \tan^{-1}\frac{f_y}{f_x}.$$

Using this value of α_1,

$$\frac{df}{dn} = f_x \frac{f_x}{\sqrt{f_x^2 + f_y^2}} + f_y \frac{f_y}{\sqrt{f_x^2 + f_y^2}}$$
$$= \sqrt{f_x^2 + f_y^2}.$$

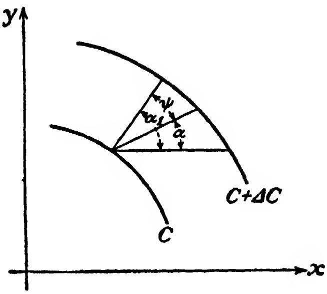

FIG. 32.

The relation (42-2) can be derived directly by use of this expression for df/dn. If α (Fig. 32) gives any direction different from the direction given by α_1, then

$$\frac{df}{ds} = f_x \cos\alpha + f_y \sin\alpha.$$

But $\alpha = \alpha_1 - \psi$, so that

$$\frac{df}{ds} = f_x(\cos\alpha_1 \cos\psi + \sin\alpha_1 \sin\psi) + f_y(\sin\alpha_1 \cos\psi - \cos\alpha_1 \sin\psi).$$

Since

$$\cos\alpha_1 = \frac{f_x}{\sqrt{f_x^2 + f_y^2}} \qquad \text{and} \qquad \sin\alpha_1 = \frac{f_y}{\sqrt{f_x^2 + f_y^2}},$$

$$\begin{aligned}\frac{df}{ds} &= f_x \frac{f_x}{\sqrt{f_x^2 + f_y^2}} \cos\psi + f_x \frac{f_y}{\sqrt{f_x^2 + f_y^2}} \sin\psi \\ &\quad + f_y \frac{f_y}{\sqrt{f_x^2 + f_y^2}} \cos\psi - f_y \frac{f_x}{\sqrt{f_x^2 + f_y^2}} \sin\psi \\ &= \frac{f_x^2 + f_y^2}{\sqrt{f_x^2 + f_y^2}} \cos\psi = \sqrt{f_x^2 + f_y^2} \cos\psi \\ &= \frac{df}{dn} \cos\psi.\end{aligned}$$

PROBLEMS

1. Find the directional derivative of $f(x, y) = x^2y + \sin xy$ at $(1, \pi/2)$, in the direction of the line making an angle of 45° with the x-axis.

2. Find

$$\frac{df}{dn} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2},$$

if $x = r\cos\theta$, $y = r\sin\theta$, and f is a function of the variables r and θ.

3. Find the directional derivative of $f(x, y) = x^3y + e^{yx}$ in the direction of the curve which, at the point (1, 1), makes an angle of 30° with the x-axis.

4. Find the normal derivative of $f(x, y) = x^2 + y^2$.

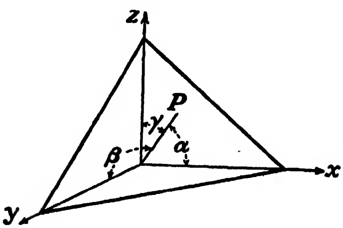

Fig. 33.

43. Tangent Plane and Normal Line to a Surface. It will be recalled that

$$Ax + By + Cz = D$$

is the equation of a plane, where the coefficients A, B, and C are called the direction components of the normal to the plane. If α, β, and γ (Fig. 33) are the direction angles made by the normal to the plane from the origin, then

$$\cos\alpha = \frac{A}{\sqrt{A^2 + B^2 + C^2}}, \qquad \cos\beta = \frac{B}{\sqrt{A^2 + B^2 + C^2}},$$
$$\cos\gamma = \frac{C}{\sqrt{A^2 + B^2 + C^2}}.$$

Therefore,

$$\cos\alpha : \cos\beta : \cos\gamma = A : B : C.$$

If the plane passes through the point (x_0, y_0, z_0), its equation can be written as

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

There is also a normal form for the equation of a plane, entirely analogous to the normal form for the equation of the straight line in the plane. This form is

$$x\cos\alpha + y\cos\beta + z\cos\gamma = p,$$

or

$$\frac{A}{\sqrt{A^2 + B^2 + C^2}}x + \frac{B}{\sqrt{A^2 + B^2 + C^2}}y + \frac{C}{\sqrt{A^2 + B^2 + C^2}}z = \frac{D}{\sqrt{A^2 + B^2 + C^2}},$$

in which $p = D/\sqrt{A^2 + B^2 + C^2}$ is the distance from the origin to the plane.

Consider a surface defined by $z = f(x, y)$, in which x and y are considered as the independent variables. Then,

$$\begin{aligned} dz &= \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy \\ &= \frac{\partial f}{\partial x}\,\Delta x + \frac{\partial f}{\partial y}\,\Delta y. \end{aligned} \tag{43-1}$$

If x_0 and y_0 are chosen, z_0 is determined by $z = f(x, y)$. Let $\Delta x = x - x_0$ and $\Delta y = y - y_0$, and denote dz by $z - z_0$. Then (43-1) becomes

$$z - z_0 = \left.\frac{\partial f}{\partial x}\right|_{(x_0,\, y_0)} (x - x_0) + \left.\frac{\partial f}{\partial y}\right|_{(x_0,\, y_0)} (y - y_0), \tag{43-2}$$

which is the equation of a plane. If this plane is cut by the plane $x = x_0$, the equation of the line of intersection is

$$z - z_0 = \left.\frac{\partial f}{\partial y}\right|_{(x_0,\, y_0)} (y - y_0),$$

and this is the tangent line to the curve $z = f(x_0, y)$ at the point (x_0, y_0, z_0). Similarly, the line of intersection of the plane defined by (43-2) and the plane $y = y_0$ is the tangent line to the curve $z = f(x, y_0)$ at (x_0, y_0, z_0). The plane defined by (43-2) is called the *tangent plane* to the surface $z = f(x, y)$ at (x_0, y_0, z_0).

The direction cosines of the normal to this plane are proportional to

$$\left.\frac{\partial f}{\partial x}\right|_{(x_0,\, y_0)}, \qquad \left.\frac{\partial f}{\partial y}\right|_{(x_0,\, y_0)}, \qquad -1.$$

The equation of the normal line to the plane (43-2) at (x_0, y_0, z_0) is therefore

$$\frac{x - x_0}{\left.\dfrac{\partial f}{\partial x}\right|_{(x_0,\, y_0)}} = \frac{y - y_0}{\left.\dfrac{\partial f}{\partial y}\right|_{(x_0,\, y_0)}} = \frac{z - z_0}{-1}. \tag{43-3}$$

This line is defined as *the normal* to the surface at (x_0, y_0, z_0). Figure 34 shows the difference between $dz = RP'$ and $\Delta z = RQ$. $P(x_0, y_0, z_0)$ is the point of tangency and $R(x_0 + \Delta x, y_0 + \Delta y, z_0)$ is in the plane $z = z_0$. PP' is the tangent plane.

In case the equation of the surface is given in the form

$$F(x, y, z) = 0,$$

the tangent plane and the normal line at (x_0, y_0, z_0) have the respective equations

$$(43\text{-}4)\qquad \left.\frac{\partial F}{\partial x}\right|_{(x_0,\,y_0,\,z_0)}(x - x_0) + \left.\frac{\partial F}{\partial y}\right|_{(x_0,\,y_0,\,z_0)}(y - y_0) + \left.\frac{\partial F}{\partial z}\right|_{(x_0,\,y_0,\,z_0)}(z - z_0) = 0$$

and

$$(43\text{-}5)\qquad \frac{x - x_0}{\left.\frac{\partial F}{\partial x}\right|_{(x_0,\,y_0,\,z_0)}} = \frac{y - y_0}{\left.\frac{\partial F}{\partial y}\right|_{(x_0,\,y_0,\,z_0)}} = \frac{z - z_0}{\left.\frac{\partial F}{\partial z}\right|_{(x_0,\,y_0,\,z_0)}}.$$

These equations follow directly from (41-2).

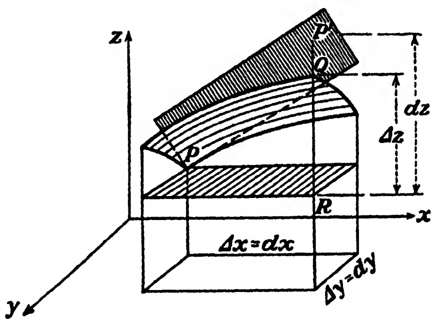

FIG. 34

Example 1. At (6, 2, 3) on the surface $x^2 + y^2 + z^2 = 49$, the tangent plane has the equation

$$2x\Big|_{(6,\,2,\,3)}(x - 6) + 2y\Big|_{(6,\,2,\,3)}(y - 2) + 2z\Big|_{(6,\,2,\,3)}(z - 3) = 0$$

or

$$6x + 2y + 3z = 49.$$

The normal line is

$$\frac{x - 6}{12} = \frac{y - 2}{4} = \frac{z - 3}{6}.$$

Example 2. For (2, 1, 4) on the surface $z = x^2 + y^2 - 1$, the tangent plane is

$$z - 4 = 2x\Big|_{(2,\,1)}(x - 2) + 2y\Big|_{(2,\,1)}(y - 1)$$

or

$$4x + 2y - z = 6.$$

The normal line is

$$\frac{x-2}{4} = \frac{y-1}{2} = \frac{z-4}{-1}.$$

PROBLEMS

1. Find the distance from the origin to the plane $x + y + z = 1$.

2. Find the equations of the tangent plane and the normal line to

(a) $2x^2 + 3y^2 + 4z^2 = 6$ at $(1, 1, \frac{1}{2})$;

(b) $\dfrac{x^2}{4} + \dfrac{y^2}{9} - \dfrac{z^2}{16} = 1$ at $(4, 3, 8)$;

(c) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ at (x_0, y_0, z_0);

(d) $x^2 + 2y^2 - z^2 = 0$ at $(1, 2, 3)$.

3. Referring to (43-4), show that

$$\cos\alpha : \cos\beta : \cos\gamma = \frac{\partial F}{\partial x} : \frac{\partial F}{\partial y} : \frac{\partial F}{\partial z},$$

where $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are direction cosines of the normal line.

4. Show that the sum of the intercepts on the coordinate axes of any tangent plane to $x^{\frac{1}{2}} + y^{\frac{1}{2}} + z^{\frac{1}{2}} = a^{\frac{1}{2}}$ is constant.

44. Space Curves. It will be recalled that a plane curve C whose equation is

$$y = f(x) \tag{44-1}$$

can be represented in infinitely many ways by a pair of parametric equations

$$\begin{aligned} x &= x(t), \\ y &= y(t) \end{aligned} \tag{44-2}$$

so chosen that when the independent variable t runs continuously through some set of values $t_1 \leq t \leq t_2$ the corresponding values of x and y, determined by (44-2), satisfy (44-1).

For example, the equation of the upper half of a unit circle with the center at the origin of the cartesian system,

$$y = \sqrt{1 - x^2},$$

can be represented parametrically as

$$\begin{aligned} x &= \cos t, \\ y &= \sin t, \end{aligned} \qquad (0 \leq t \leq \pi),$$

or

$$x = t,$$
$$y = \sqrt{1 - t^2}, \qquad (0 \le t \le 1),$$

or

$$x = 2t,$$
$$y = \sqrt{1 - 4t^2}, \qquad (0 \le t \le \tfrac{1}{2}).$$

Similarly, a space curve C can be represented by means of a set of equations

$$\text{(44-3)} \qquad \begin{cases} x = x(t), \\ y = y(t), \\ z = z(t) \end{cases}$$

so selected that when t runs through some set of values the coordinates of the point $P(x, y, z)$, defined by (44-3), trace out the desired curve C.

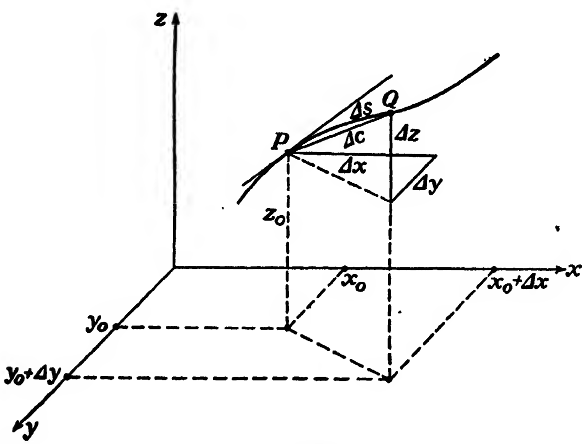

FIG. 35.

It will be assumed that the functions in (44-2) and (44-3) possess continuous derivatives with respect to t, which implies that the curve C has a continuously turning tangent as the point P moves along the curve.

Let $P(x_0, y_0, z_0)$ (Fig. 35) be a point of the curve C defined by (44-3) that corresponds to some value t_0 of the parameter t, and let Q be the point $(x_0 + \Delta x,\ y_0 + \Delta y,\ z_0 + \Delta z)$ that corresponds to $t = t_0 + \Delta t$. The direction ratios of the line PQ

joining P and Q are

$$\frac{\Delta x}{\Delta c}:\frac{\Delta y}{\Delta c}:\frac{\Delta z}{\Delta c} = \frac{\Delta x}{\Delta t}:\frac{\Delta y}{\Delta t}:\frac{\Delta z}{\Delta t}.$$

If Δt is allowed to approach zero, Δx, Δy, and Δz all tend to zero, so that the direction ratios of the tangent line at $P(x_0, y_0, z_0)$ are proportional to $(dx/dt)_{t=t_0}:(dy/dt)_{t=t_0}:(dz/dt)_{t=t_0}$. Hence, the equation of the tangent line to C at P is

$$\frac{x - x_0}{x'(t_0)} = \frac{y - y_0}{y'(t_0)} = \frac{z - z_0}{z'(t_0)},$$

where primes denote derivatives with respect to t.

Example. The equation of the tangent line to the circular helix

$$x = a \cos t,$$
$$y = a \sin t,$$
$$z = at,$$

at $t = \pi/6$, is

$$\frac{x - \frac{\sqrt{3}}{2}a}{-\frac{a}{2}} = \frac{y - \frac{a}{2}}{\frac{\sqrt{3}}{2}a} = \frac{z - \frac{\pi a}{6}}{a}.$$

The element of arc ds is given by

$$(ds)^2 = (dx)^2 + (dy)^2 + (dz)^2,$$

so that the length of a space curve C can be calculated from

$$L = \int_{t_1}^{t_2} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt.$$

The length of the part of the helix between the points $(a, 0, 0)$ and $(0, a, \pi a/2)$ is

$$L = \frac{\sqrt{2}}{2}\pi a.$$

45. Directional Derivatives in Space. There is no essential difficulty in extending the results of Sec. 42 to any number of variables. Thus, if $u = f(x, y, z)$ is a suitably restricted function of the independent variables x, y, and z, then the directional derivative along a space curve whose tangent line at some point $P(x, y, z)$ (Fig. 35) has the direction cosines $\cos(x, s)$, $\cos(y, s)$,

and $\cos(z, s)$ is

$$\frac{du}{ds} = \frac{\partial u}{\partial x}\cos(x, s) + \frac{\partial u}{\partial y}\cos(y, s) + \frac{\partial u}{\partial z}\cos(z, s)$$
$$= \frac{\partial u}{\partial x}\frac{dx}{ds} + \frac{\partial u}{\partial y}\frac{dy}{ds} + \frac{\partial u}{\partial z}\frac{dz}{ds}.$$

The magnitude of the normal derivative to the surface $u =$ const. is given by

$$\frac{du}{dn} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2}. \tag{45-1}$$

The vector that is normal to the surface $u =$ const. and whose magnitude is du/dn is called the *gradient* of u.

PROBLEM

1. Find the equation of the tangent line to the helix

$$x = a\cos t, \qquad y = a\sin t, \qquad z = at,$$

at the point where $t = \pi/4$. Find the length of the helix between the points $t = 0$ and $t = \pi/4$.

2. Find the directional derivative of $f = xyz$ at $(1, 2, 3)$ in the direction of the line that makes equal angles with the coordinate axes.

3. Find the normal derivative of $f = x^2 + y^2 + z^2$ at $(1, 2, 3)$.

4. Show that the square root of the sum of the squares of the directional derivatives in three perpendicular directions is equal to the normal derivative.

5. Express the normal derivative (45-1) in spherical and cylindrical coordinates, for which the equations of transformation are

(a) $x = r\sin\theta\cos\varphi$, $y = r\sin\theta\sin\varphi$, $z = r\cos\theta$;
(b) $x = r\sin\theta$, $y = r\cos\theta$, $z = z$.

6. What is the direction of the curve $x = t$, $y = t^2$, $z = t^3$ at the point $(1, 1, 1)$?

7. Show that the condition that the surfaces $f(x, y, z) = 0$ and $g(x, y, z) = 0$ intersect orthogonally is that

$$\frac{\partial f}{\partial x}\frac{\partial \varphi}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial \varphi}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial \varphi}{\partial z} = 0.$$

8. Show that the surfaces

$$xyz = 1 \qquad \text{and} \qquad \frac{x^2}{2} + \frac{y^2}{2} - \frac{z^2}{1} = 1$$

intersect at right angles.

9. Find the angle between the normals to the tangent planes to the surfaces $x^2 + y^2 + z^2 = 6$ and $2x^2 + 3y^2 + z^2 = 9$ at the point (1, 1, 2).

10. Show that the direction of the tangent line to the curve of intersection of the surfaces $f(x, y, z) = 0$ and $g(x, y, z) = 0$ is given by

$$\cos\alpha : \cos\beta : \cos\gamma = \begin{vmatrix} f_y & f_z \\ g_y & g_z \end{vmatrix} : \begin{vmatrix} f_z & f_x \\ g_z & g_x \end{vmatrix} : \begin{vmatrix} f_x & f_y \\ g_x & g_y \end{vmatrix}.$$

Hint: Let (x_0, y_0, z_0) be a point on the curve of intersection, and find the line of intersection of the tangent planes to the surfaces at the point (x_0, y_0, z_0).

46. Higher Partial Derivatives. The partial derivatives f_{x_1}, $f_{x_2}, \cdots, f_{x_n}$ of $f(x_1, x_2, \cdots, x_n)$ are functions of $x_1, x_2, \cdots, x_n$ and may have partial derivatives with respect to some or all of these variables. These derivatives are called second partial derivatives of $f(x_1, x_2, \cdots, x_n)$. If there are only two independent variables x and y, then $f(x, y)$ may have the second partial derivatives

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \equiv \frac{\partial^2 f}{\partial x^2} \equiv f_{xx},$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \equiv \frac{\partial^2 f}{\partial y\,\partial x} \equiv f_{xy},$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \equiv \frac{\partial^2 f}{\partial x\,\partial y} \equiv f_{yx},$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \equiv \frac{\partial^2 f}{\partial y^2} \equiv f_{yy}.$$

It should be noticed that f_{xy} means that $\partial f/\partial x$ is first found and then $\dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right)$ is determined, so that the subscripts indicate the order in which the derivatives are taken. In

$$\frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

the order is in keeping with the meaning of the symbol, so that the order appears as the reverse of the order in which the derivatives are taken.

It can be proved* that, if f_{xy} and f_{yx} are continuous functions of x and y, then $f_{xy} = f_{yx}$, so that the order of differentiation is

* See Sokolnikoff, I. S., Advanced Calculus, Sec. 31.

immaterial. Similarly, when third partial derivatives are found, $f_{xyx} = f_{xxy} = f_{yxx}$ and $f_{xyy} = f_{yxy} = f_{yyx}$, if these derivatives are continuous.

Example. If $f(x, y) = e^{xy}$, then

$$f_x = ye^{xy}, \qquad f_y = xe^{xy}, \qquad f_{xx} = y^2e^{xy},$$
$$f_{xy} = f_{yx} = e^{xy}(xy + 1), \qquad f_{yy} = x^2e^{xy}.$$

PROBLEMS

1. Verify that $\dfrac{\partial^2 f}{\partial x\,\partial y} = \dfrac{\partial^2 f}{\partial y\,\partial x}$ for

(*a*) $f = \cos xy^2$, (*b*) $f = \sin^2 x \cos y$, (*c*) $f = e^{y/x}$.

2. Prove that if

(*a*) $f(x, y) = \log(x^2 + y^2) + \tan^{-1}\dfrac{y}{x}$, then $\dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} = 0$;

(*b*) $f(x, y, z) = (x^2 + y^2 + z^2)^{-1/2}$, then $\dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} + \dfrac{\partial^2 f}{\partial z^2} = 0$.

3. If $u = x^2 + y^2$ and $\begin{cases} x = s + 3t, \\ y = 2s - t, \end{cases}$ find $\dfrac{\partial^2 u}{\partial s^2}$ and $\dfrac{\partial^2 u}{\partial t^2}$.

4. If $u = f(x, y)$ and $\begin{cases} x = r\cos\theta, \\ y = r\sin\theta, \end{cases}$ find $\dfrac{\partial^2 u}{\partial r^2}$ and $\dfrac{\partial^2 u}{\partial \theta^2}$.

5. Use the results obtained in Prob. 6(*b*), Sec. 41, in order to show that $\dfrac{\partial^2 u}{\partial x^2} = u(3v^2 - u^2)/(u^2 + v^2)^3$. Find $\dfrac{\partial^2 u}{\partial y\,\partial x}$ and $\dfrac{\partial^2 u}{\partial y^2}$.

6. If $w = w(x, y)$, where $x = x(u, v)$, $y = y(u, v)$, $\partial x/\partial u = \partial y/\partial v$, and $\partial x/\partial v = -\partial y/\partial u$, show that

$$\frac{\partial^2 w}{\partial u^2} + \frac{\partial^2 w}{\partial v^2} = \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right)\left[\left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial x}{\partial v}\right)^2\right].$$

7. Show that the expressions

$$V_1 = \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 \qquad \text{and} \qquad V_2 = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}$$

upon change of variable by means of $x = r\cos\theta$ and $y = r\sin\theta$ become

$$V_1 = \left(\frac{\partial z}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2 \qquad \text{and} \qquad V_2 = \frac{\partial^2 z}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2} + \frac{1}{r}\frac{\partial z}{\partial r}.$$

8. If $V = f(x + ct) + g(x - ct)$, where f and g are any functions possessing continuous second derivatives, show that

$$\frac{\partial^2 V}{\partial t^2} = c^2\frac{\partial^2 V}{\partial x^2}.$$

9. Show that if $x = e^r \cos\theta$ and $y = e^r \sin\theta$, then

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = e^{-2r}\left(\frac{\partial^2 V}{\partial r^2} + \frac{\partial^2 V}{\partial \theta^2}\right).$$

10. If $V_1(x, y, z)$ and $V_2(x, y, z)$ satisfy the equation

$$\nabla^2 V \equiv \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0,$$

show that

$$U \equiv V_1(x, y, z) + (x^2 + y^2 + z^2)V_2(x, y, z)$$

satisfies the equation

$$\nabla^2\nabla^2 U = 0, \qquad \text{where} \qquad \nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

47. Taylor's Series for Functions of Two Variables. This section contains a formal development of a function of two variables, $f(x, y)$, in a series analogous to the Taylor's series development of a function of a single variable. It is assumed that the series obtained here converges to the value of the function $f(x, y)$, but the analysis of the conditions under which this convergence will occur is too involved to be discussed in this book.

Consider $f(x, y)$, which is a function of the two variables x and y, and let it have continuous partial derivatives of all orders. Let

$$x = a + \alpha t \qquad \text{and} \qquad y = b + \beta t, \tag{47-1}$$

where a, b, α, and β are constants and t is a variable. Then

$$f(x, y) = f(a + \alpha t, b + \beta t) \equiv F(t). \tag{47-2}$$

If $F(t)$ is expanded in Maclaurin's series, there results

$$F(t) = F(0) + F'(0)t + \frac{F''(0)}{2!}t^2 + \frac{F'''(0)}{3!}t^3 + \cdots. \tag{47-3}$$

From (47-2) and (47-1), it follows that

$$\begin{aligned} F'(t) &= f_x(x, y)\frac{dx}{dt} + f_y(x, y)\frac{dy}{dt} \\ &= f_x(x, y)\alpha + f_y(x, y)\beta. \end{aligned}$$

Then,

$$\begin{aligned} F''(t) &= [f_{xx}(x, y)\alpha + f_{yx}(x, y)\beta]\frac{dx}{dt} + [f_{xy}(x, y)\alpha + f_{yy}(x, y)\beta]\frac{dy}{dt} \\ &= f_{xx}(x, y)\alpha^2 + 2f_{xy}(x, y)\alpha\beta + f_{yy}(x, y)\beta^2, \end{aligned}$$

and

$$F'''(t) = [f_{xxx}(x, y)\alpha^2 + 2f_{xyx}(x, y)\alpha\beta + f_{yyx}(x, y)\beta^2]\frac{dx}{dt}$$
$$+ [f_{xxy}(x, y)\alpha^2 + 2f_{xyy}(x, y)\alpha\beta + f_{yyy}(x, y)\beta^2]\frac{dy}{dt}$$
$$= f_{xxx}(x, y)\alpha^3 + 3f_{xxy}(x, y)\alpha^2\beta + 3f_{xyy}(x, y)\alpha\beta^2 + f_{yyy}(x, y)\beta^3.$$

Higher order derivatives of $F(t)$ can be obtained by continuing this process, but the form is evident from those already obtained. Symbolically expressed,

$$F'(t) = \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)f(x, y) \equiv \alpha\frac{\partial f}{\partial x} + \beta\frac{\partial f}{\partial y},$$
$$F''(t) = \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)^2 f(x, y) \equiv \alpha^2\frac{\partial^2 f}{\partial x^2} + 2\alpha\beta\frac{\partial^2 f}{\partial x\,\partial y} + \beta^2\frac{\partial^2 f}{\partial y^2},$$
$$F'''(t) = \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)^3 f(x, y) \equiv \alpha^3\frac{\partial^3 f}{\partial x^3} + 3\alpha^2\beta\frac{\partial^3 f}{\partial x^2\,\partial y} + 3\alpha\beta^2\frac{\partial^3 f}{\partial x\,\partial y^2} + \beta^3\frac{\partial^3 f}{\partial y^3}.$$

Then,

$$F^{(n)}(t) = \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)^n f(x, y) \equiv \alpha^n\frac{\partial^n f}{\partial x^n} + C_1^n\alpha^{n-1}\beta\frac{\partial^n f}{\partial x^{n-1}\,\partial y} + \cdots + C_{n-1}^n\alpha\beta^{n-1}\frac{\partial^n f}{\partial x\,\partial y^{n-1}} + \beta^n\frac{\partial^n f}{\partial y^n},$$

where

$$C_r^n \equiv \frac{n!}{r!(n-r)!}.$$

Since $t = 0$ gives $x = a$ and $y = b$, it follows that

$$F(0) = f(a, b), \qquad F'(0) = \alpha f_x(a, b) + \beta f_y(a, b), \cdots.$$

Substituting these expressions in (47-3) gives

$$F(t) \equiv f(x, y) = f(a, b) + [\alpha f_x(a, b) + \beta f_y(a, b)]t + [\alpha^2 f_{xx}(a, b) + 2\alpha\beta f_{xy}(a, b) + \beta^2 f_{yy}(a, b)]\frac{t^2}{2!} + \cdots.$$

Since $\alpha t = x - a$ and $\beta t = y - b$, the expansion becomes

$$(47\text{-}4)\quad f(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) + \frac{1}{2!}[f_{xx}(a, b)(x - a)^2 + 2f_{xy}(a, b)(x - a)(y - b) + f_{yy}(a, b)(y - b)^2] + \cdots.$$

This is Taylor's expansion for a function $f(x, y)$ about the point (a, b).

Another form that is frequently used is obtained by replacing $(x - a)$ by h and $(y - b)$ by k, so that $x = a + h$ and $y = b + k$. Then,

$$(47\text{-}5)\quad f(a + h, b + k) = f(a, b) + f_x(a, b)h + f_y(a, b)k + \frac{1}{2!}[f_{xx}(a, b)h^2 + 2f_{xy}(a, b)hk + f_{yy}(a, b)k^2] + \cdots.$$

This formula is frequently written symbolically as

$$f(a + h, b + k) = f(a, b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)f(a, b) + \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b) + \cdots.$$

Example. Obtain the expansion of $\tan^{-1}\frac{y}{x}$ about $(1, 1)$ up to the third-degree terms. Here, $f(x, y) = \tan^{-1}\frac{y}{x}$, so that

$$\begin{aligned}
f(x, y) &= \tan^{-1}\frac{y}{x}, & f(1, 1) &= \tan^{-1} 1 = \frac{\pi}{4};\\
f_x(x, y) &= -\frac{y}{x^2 + y^2}, & f_x(1, 1) &= -\frac{1}{2};\\
f_y(x, y) &= \frac{x}{x^2 + y^2}, & f_y(1, 1) &= \frac{1}{2};\\
f_{xx}(x, y) &= \frac{2xy}{(x^2 + y^2)^2}, & f_{xx}(1, 1) &= \frac{1}{2};\\
f_{xy}(x, y) &= \frac{y^2 - x^2}{(x^2 + y^2)^2}, & f_{xy}(1, 1) &= 0;\\
f_{yy}(x, y) &= \frac{-2xy}{(x^2 + y^2)^2}, & f_{yy}(1, 1) &= -\frac{1}{2}.
\end{aligned}$$

Then,

$$\tan^{-1}\frac{y}{x} = \frac{\pi}{4} - \frac{1}{2}(x - 1) + \frac{1}{2}(y - 1) + \frac{1}{2!}\left[\frac{1}{2}(x - 1)^2 - \frac{1}{2}(y - 1)^2\right] + \cdots.$$

PROBLEMS

1. Obtain the expansion for $xy^2 + \cos xy$ about $(1, \pi/2)$ up to the third-degree terms.

2. Expand $f(x, y) = e^{xy}$ at $(1, 1)$, obtaining three terms.
3. Expand $e^x \cos y$ at $(0, 0)$ up to the fourth-degree terms.
4. Show that, for small values of x and y,

$$e^x \sin y = y + xy \text{ (approx.)},$$

and

$$e^x \log (1 + y) = y + xy - \frac{y^2}{2} \text{ (approx.)}.$$

5. Expand $f(x, y) = x^3y + x^2y + 1$ about $(0, 1)$.
6. Expand $(1 - x^2 - y^2)^{1/2}$ about $(0, 0)$ up to the third-degree terms.
7. Show that the development obtained in Prob. 6 is identical with the binomial expansion of $[1 - (x^2 + y^2)]^{1/2}$.

48. Maxima and Minima of Functions of One Variable. A function $f(x)$ is said to have a *maximum* at $x = a$, if

$$\Delta^+ \equiv f(a + h) - f(a) < 0,$$

and

$$\Delta^- \equiv f(a - h) - f(a) < 0,$$

for all sufficiently small positive values of h. If Δ^+ and Δ^- are both positive for all small positive values of h, then $f(x)$ is said to have a *minimum* at $x = a$.

It is shown in the elementary calculus that, if the function $f(x)$ has a derivative at $x = a$, then the necessary condition for a maximum or a minimum is the vanishing of $f'(x)$ at the point $x = a$. Of course, the function $f(x)$ may attain a maximum or a minimum at $x = a$ without having $f'(a) = 0$, but this can occur only if $f'(x)$ ceases to exist at the critical point (see Fig. 36).

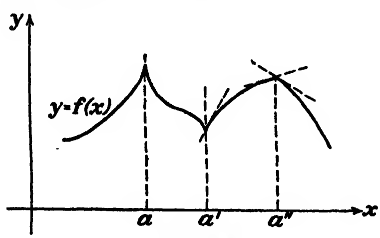

FIG. 36.

Let it be supposed that $f(x)$ has a continuous derivative of order n in some interval about the point $x = a$. Then it follows from Taylor's formula that

$$\Delta^+ \equiv f(a + h) - f(a)$$
$$= f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}h^{n-1} + \frac{f^{(n)}(a + \theta_1 h)}{n!}h^n,$$

where $0 < \theta_1 < 1$, and

$$\begin{aligned}\Delta^- &\equiv f(a-h) - f(a) \\ &= -f'(a)h + \frac{f''(a)}{2!}h^2 - \cdots + (-1)^{n-1}\frac{f^{(n-1)}(a)}{(n-1)!}h^{n-1} \\ &\qquad + (-1)^n \frac{f^{(n)}(a-\theta_2 h)}{n!}h^n,\end{aligned}$$

where $0 < \theta_2 < 1$. Let it be assumed further that the first $n - 1$ derivatives of $f(x)$ vanish at $x = a$ but that $f^{(n)}(a)$ is not zero. Then

$$\Delta^+ = \frac{f^{(n)}(a + \theta_1 h)}{n!}h^n$$

and

$$\Delta^- = (-1)^n \frac{f^{(n)}(a - \theta_2 h)}{n!}h^n.$$

Since $f^{(n)}(x)$ is assumed to be continuous in some interval about the point $x = a$, $f^{(n)}(a + \theta_1 h)$ and $f^{(n)}(a - \theta_2 h)$ will have the same sign for sufficiently small values of h. Consequently, the signs of Δ^+ and Δ^- will be opposite unless n is an even number. But if $f(x)$ is to have a maximum or a minimum at $x = a$, then Δ^+ and Δ^- must be of the same sign. Accordingly, the necessary condition for a maximum or a minimum of $f(x)$ at $x = a$ is that the first non-vanishing derivative of $f(x)$, at $x = a$, be of even order. Moreover, since both Δ^+ and Δ^- are negative if $f(x)$ is a maximum, it follows that $f^{(n)}(a)$ must be negative. A similar argument shows that, if $f(x)$ has a minimum at $x = a$, then the first non-vanishing derivative of $f(x)$ at $x = a$ must be of even order and positive.

If the first non-vanishing derivative of $f(x)$ at $x = a$ is of odd order and $f''(a) = 0$, then the point $x = a$ is called a *point of inflection*.

Example. Investigate $f(x) = x^5 - 5x^4$ for maxima and minima. Now,

$$f'(x) = 5x^4 - 20x^3,$$

which is zero when $x = 0$ and $x = 4$. Then,

$$\begin{aligned} f''(x) &= 20x^3 - 60x^2, \quad f''(0) = 0, \quad f''(4) = 320; \\ f'''(x) &= 60x^2 - 120x, \quad f'''(0) = 0; \\ f^{\mathrm{IV}}(x) &= 120x - 120, \quad f^{\mathrm{IV}}(0) = -120. \end{aligned}$$

Since $f''(4) > 0$, $f(4) = -256$ is a minimum; and since $f^{\mathrm{IV}}(0) < 0$, $f(0) = 0$ is a maximum.

PROBLEMS

1. Examine the following for maxima and minima:

(a) $y = x^4 - 4x^3 + 1$;
(b) $y = x^3(x - 5)^2$;
(c) $y = x + \cos x$.

2. Find the minimum of the function $y = x^x$, where $x > 0$.
Hint: Consider the minimum of $\log y$.

3. Show that $x = 0$ gives the minimum value of the function

$$y = e^x + e^{-x} + 2 \cos x.$$

4. Find maxima, minima, and points of inflection, and sketch the curves, for the following:

(a) $y = 3x + 4 \sin x + \sin 2x$;
(b) $y = 3x - 4 \sin x + \sin 2x$;
(c) $y = 6x + 8 \sin x + \sin 2x$.

5. Find the maximum and minimum values of the function

$$y = x \sin x + 2 \cos x.$$

6. Find maxima, minima, and points of inflection, and sketch the curves, for the following:

(a) $y = x \log x$;
(b) $x^5 - (y - x^2)^2 = 0$.

49. Maxima and Minima of Functions of Several Variables. A function of two variables $f(x, y)$ is said to have a maximum at (a, b), if $f(a + h, b + k) - f(a, b) < 0$ for sufficiently small positive and negative values of h and k, and a minimum, if $f(a + h, b + k) - f(a, b) > 0$.

Geometrically, this means that when the point (a, b, c) on the surface $z = f(x, y)$ is higher than all neighboring points, then (a, b, c) is a maximum; and when (a, b, c) is lower than all neighboring points, it is a minimum point. At a maximum or a minimum point (a, b, c) the curves in which the planes $x = a$ and $y = b$ cut the surface have maxima or minima. Therefore, $f_x(a, b) = 0$ and $f_y(a, b) = 0$. The conditions $f_x = 0$ and $f_y = 0$ can be solved simultaneously to give the critical values.

The testing of the critical values for maxima and minima is more difficult than in the case of functions of one variable. However in many applied problems the physical interpretation

will determine whether or not the critical values yield maxima or minima or neither. An analytical criterion can be established for the case of two variables in a manner analogous to the method used for one variable. By the use of Taylor's expansion, it can be shown that if $f_x(a, b) = 0$ and $f_y(a, b) = 0$, then $f(a, b)$ is a maximum if

$$D \equiv f_{xy}^2(a, b) - f_{xx}(a, b)f_{yy}(a, b) < 0$$

with

$$f_{xx}(a, b) < 0 \qquad \text{and} \qquad f_{yy}(a, b) < 0,$$

and a minimum if

$$D \equiv f_{xy}^2(a, b) - f_{xx}(a, b)f_{yy}(a, b) < 0$$

with

$$f_{xx}(a, b) > 0 \qquad \text{and} \qquad f_{yy}(a, b) > 0.$$

In case

$$f_{xy}^2(a, b) - f_{xx}(a, b)f_{yy}(a, b) > 0,$$

$f(a, b)$ is neither a maximum nor a minimum. If

$$f_{xy}^2(a, b) - f_{xx}(a, b)f_{yy}(a, b) = 0,$$

the test gives no information, just as $f''(x) = 0$ gives no criterion in the case of one variable.

These considerations can be extended to functions of more than two variables. Thus, in the case of a function $f(x, y, z)$ of three variables,

$$\frac{\partial f}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} = 0, \qquad \frac{\partial f}{\partial z} = 0$$

is the necessary condition for a maximum or a minimum.

Example 1. A long piece of tin 12 in. wide is made into a trough by bending up the sides to form equal angles with the base (Fig. 37). Find the amount to be bent up and the angle of inclination of the sides that will make the carrying capacity a maximum.

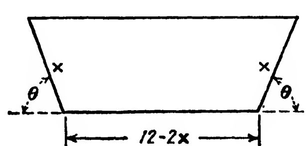

Fig. 37.

The volume will be a maximum if the area of the trapezoidal cross section is a maximum. The area is

$$A = 12x \sin\theta - 2x^2 \sin\theta + x^2 \sin\theta \cos\theta;$$

for $12 - 2x$ is the lower base, $12 - 2x + 2x \cos\theta$ is the upper base,

and $x \sin \theta$ is the altitude. Then,

$$\frac{\partial A}{\partial \theta} = 12x \cos \theta - 2x^2 \cos \theta + x^2 \cos^2 \theta - x^2 \sin^2 \theta$$
$$= x(12 \cos \theta - 2x \cos \theta + x \cos^2 \theta - x \sin^2 \theta)$$

and

$$\frac{\partial A}{\partial x} = 2 \sin \theta(6 - 2x + x \cos \theta).$$

$\partial A/\partial x = 0$ and $\partial A/\partial \theta = 0$, if $\sin \theta = 0$ and $x = 0$, which, from physical considerations, cannot give a maximum.

There remain to be satisfied

$$6 - 2x + x \cos \theta = 0$$

and

$$12 \cos \theta - 2x \cos \theta + x \cos^2 \theta - x \sin^2 \theta = 0.$$

Solving the first equation for x and substituting in the second yield, upon simplification,

$$\cos \theta = \tfrac{1}{2} \qquad \text{or} \qquad \theta = 60°, \qquad \text{and} \qquad x = 4.$$

Since physical considerations show that a maximum exists, $x = 4$ and $\theta = 60°$ must give the maximum.

Example 2. Find the maxima and minima of the surface

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2cz.$$

Now,

$$\frac{\partial z}{\partial x} = \frac{1}{c}\frac{x}{a^2}, \qquad \frac{\partial z}{\partial y} = -\frac{1}{c}\frac{y}{b^2},$$

which vanish when $x = y = 0$. But

$$\frac{\partial^2 z}{\partial x^2} = \frac{1}{a^2c}, \qquad \frac{\partial^2 z}{\partial y^2} = -\frac{1}{b^2c}, \qquad \frac{\partial^2 z}{\partial x\, \partial y} = 0.$$

Hence, $D = 1/a^2b^2c^2$ and, consequently, there is no maximum or minimum at $x = y = 0$. The surface under consideration is a saddle-shaped surface called a *hyperbolic paraboloid*. The points for which the first partial derivatives vanish and $D > 0$ are called *minimax*. The reason for this odd name appears from a consideration of the shape of the hyperbolic paraboloid near the origin of the coordinate system. The reader will benefit from sketching it in the vicinity of (0, 0, 0).

PROBLEMS

1. Divide a into three parts such that their product is a maximum. Test by using the second derivative criterion.

2. Find the volume of the largest rectangular parallelepiped that can be inscribed in the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

3. Find the dimensions of the largest rectangular parallelepiped that has three faces in the coordinate planes and one vertex in the plane

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1.$$

4. A pentagonal frame is composed of a rectangle surmounted by an isosceles triangle. What are the dimensions for maximum area of the pentagon if the perimeter is given as P?

5. A floating anchorage is designed with a body in the form of a right-circular cylinder with equal ends that are right-circular cones. If the volume is given, find the dimensions giving the minimum surface area.

6. Given n points P_i whose coordinates are (x_i, y_i, z_i), $(i = 1, 2, \cdots, n)$. Show that the coordinates of the point $P(x, y, z)$, such that the sum of the squares of the distances from P to the P_i is a minimum, are given by

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i, \frac{1}{n}\sum_{i=1}^{n} y_i, \frac{1}{n}\sum_{i=1}^{n} z_i\right).$$

50. Constrained Maxima and Minima. In a large number of practical and theoretical investigations, it is required that a maximum or minimum value of a function be found when the variables are connected by some relation. Thus, it may be required to find a maximum of $u = f(x, y, z)$, where x, y, and z are connected by the relation $\varphi(x, y, z) = 0$. The resulting maximum is called a *constrained maximum.*

The method of obtaining maxima and minima described in the preceding section can be used to solve a problem of constrained maxima and minima, as follows: If the constraining relation $\varphi(x, y, z) = 0$ can be solved for one of the variables, say z, in terms of the remaining two variables, and if the resulting expression is substituted for z in $u = f(x, y, z)$, there will be obtained a function $u = F(x, y)$. The values of x and y that yield maxima and minima of u can be found by the methods of Sec. 49. However, the solution of $\varphi(x, y, z) = 0$ for any one of the variables may be extremely difficult, and it is desirable to consider an ingenious device used by Lagrange.

To avoid circumlocution the maximum and minimum values of a function of any number of variables will be called its *extremal values.* It follows from Sec. 49 that the necessary condition for the existence of an extremum of a differentiable function $f(x_1, x_2, \cdots, x_n)$ is the vanishing of the first partial derivatives of the function with respect to the independent variables $x_1, x_2, \cdots, x_n$. Inasmuch as the differential of a function is defined as

$$df \equiv \frac{\partial f}{\partial x_1}\,dx_1 + \frac{\partial f}{\partial x_2}\,dx_2 + \cdots + \frac{\partial f}{\partial x_n}\,dx_n,$$

it is clear that df vanishes for those values of $x_1, x_2, \cdots, x_n$ for which the function has extremal values. Conversely, since the variables x_i are assumed to be independent, the vanishing of the differential is the necessary condition for an extremum.

It is not difficult to see that, even when some of the variables are not independent, the vanishing of the total differential is the necessary condition for an extremum. Thus, consider a function

$$u = f(x, y, z), \tag{50-1}$$

where one of the variables, say z, is connected with x and y by some constraining relation

$$\varphi(x, y, z) = 0. \tag{50-2}$$

Regarding x and y as the independent variables, the necessary conditions for an extremum give $\partial u/\partial x = 0$ and $\partial u/\partial y = 0$, or

$$\frac{\partial u}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 0,$$
$$\frac{\partial u}{\partial y} = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} = 0.$$

Then the total differential

$$\frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\left(\frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy\right) = 0;$$

and since the expression in the parenthesis is precisely dz, it follows that

$$\frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz = 0. \tag{50-3}$$

The total differential of the constraining relation (50-2) is

$$\frac{\partial \varphi}{\partial x}\,dx + \frac{\partial \varphi}{\partial y}\,dy + \frac{\partial \varphi}{\partial z}\,dz = 0. \tag{50-4}$$

Let this equation be multiplied by some undetermined multiplier λ and then added to (50-3). The result is

$$\left(\frac{\partial f}{\partial x} + \lambda\frac{\partial \varphi}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y} + \lambda\frac{\partial \varphi}{\partial y}\right)dy + \left(\frac{\partial f}{\partial z} + \lambda\frac{\partial \varphi}{\partial z}\right)dz = 0.$$

Now, if λ is so chosen that

$$\begin{cases} \dfrac{\partial f}{\partial x} + \lambda\dfrac{\partial \varphi}{\partial x} = 0, \\ \dfrac{\partial f}{\partial y} + \lambda\dfrac{\partial \varphi}{\partial y} = 0, \\ \dfrac{\partial f}{\partial z} + \lambda\dfrac{\partial \varphi}{\partial z} = 0, \\ \varphi(x, y, z) = 0, \end{cases} \tag{50-5}$$

then the necessary condition for an extremum of (50-1) will surely be satisfied.

Thus, in order to determine the extremal values of (50-1), all that is necessary is to obtain the solution of the system of Eqs. (50-5) for the four unknowns x, y, z, and λ. The multiplier λ is called a *Lagrangian multiplier.*

Example 1. Find the maximum and the minimum distances from the origin to the curve

$$5x^2 + 6xy + 5y^2 - 8 = 0.$$

The problem here is to determine the extremal values of

$$f(x, y) = x^2 + y^2$$

subject to the condition

$$\varphi(x, y) \equiv 5x^2 + 6xy + 5y^2 - 8 = 0.$$

Equations (50-5) in this case become

$$\begin{aligned} 2x + \lambda(10x + 6y) &= 0, \\ 2y + \lambda(6x + 10y) &= 0, \\ 5x^2 + 6xy + 5y^2 - 8 &= 0. \end{aligned}$$

Multiplying the first of these equations by y and the second by x and

then subtracting give

$$6\lambda(y^2 - x^2) = 0,$$

so that $y = \pm x$. Substituting these values of y in the third equation gives two equations for the determination of x, namely,

$$2x^2 = 1 \qquad \text{and} \qquad x^2 = 2.$$

The first of these gives $f \equiv x^2 + y^2 = 1$, and the second gives $f \equiv x^2 + y^2 = 4$. Obviously, the first value is a minimum, whereas the second is a maximum. The curve is an ellipse of semiaxes 2 and 1 whose major axis makes an angle of 45° with the x-axis.

Example 2. Find the dimensions of the rectangular box, without a top, of maximum capacity whose surface is 108 sq. in.

The function to be maximized is

$$f(x, y, z) \equiv xyz,$$

subject to the condition

$$xy + 2xz + 2yz = 108. \tag{50-6}$$

The first three of Eqs. (50-5) become

$$\begin{cases} yz + \lambda(y + 2z) = 0, \\ xz + \lambda(x + 2z) = 0, \\ xy + \lambda(2x + 2y) = 0. \end{cases} \tag{50-7}$$

In order to solve these equations, multiply the first by x, the second by y, and the last by z, and add. There results

$$\lambda(2xy + 4xz + 4yz) + 3xyz = 0,$$

or

$$\lambda(xy + 2xz + 2yz) + \tfrac{3}{2}xyz = 0.$$

Substituting from (50-6) gives

$$108\lambda + \tfrac{3}{2}xyz = 0,$$

or

$$\lambda = -\frac{xyz}{72}.$$

Substituting this value of λ in (50-7) and dividing out common factors give

$$1 - \frac{x}{72}(y + 2z) = 0,$$

$$1 - \frac{y}{72}(x + 2z) = 0,$$

$$1 - \frac{z}{72}(2x + 2y) = 0.$$

From the first two of these equations, it is evident that $x = y$. The substitution of $x = y$ in the third equation gives $z = 18/y$. Substituting for y and z in the first equation yields $x = 6$. Thus, $x = 6$, $y = 6$, and $z = 3$ give the desired dimensions.

PROBLEMS

1. Work Probs. 1, 2, and 3, Sec. 49, by using Lagrangian multipliers.

2. Prove that the point of intersection of the medians of a triangle possesses the property that the sum of the squares of its distances from the vertices is a minimum.

3. Find the maximum and the minimum of the sum of the angles made by a line from the origin with (*a*) the coordinate axes of a cartesian system; (*b*) the coordinate planes.

4. Find the maximum distance from the origin to the folium of Descartes $x^3 + y^3 - 3axy = 0$.

5. Find the shortest distance from the origin to the plane

$$ax + by + cz = d.$$

51. Differentiation under the Integral Sign. Integrals whose integrands contain a parameter have already occurred in the first chapter. Thus, the length of arc of an ellipse is expressible as a definite integral containing the eccentricity of the ellipse as a parameter.*

Consider a definite integral

$$\varphi(\alpha) = \int_{u_0}^{u_1} f(x, \alpha)\, dx, \tag{51-1}$$

in which the integrand contains a parameter α and where u_0 and u_1 are constants. As a specific illustration, let

$$\varphi(\alpha) = \int_0^{\frac{\pi}{2}} \sin \alpha x\, dx.$$

In this case the indefinite integral

$$F(x, \alpha) = \int \sin \alpha x\, dx = -\frac{\cos \alpha x}{\alpha} + C$$

is a function of both x and α; but, upon substitution of the limits, there appears a function of α alone, namely

$$\varphi(\alpha) = \int_0^{\frac{\pi}{2}} \sin \alpha x\, dx = -\left.\frac{\cos \alpha x}{\alpha}\right|_0^{\frac{\pi}{2}} = \frac{1}{\alpha}\left(1 - \cos \frac{\pi\alpha}{2}\right).$$

* See Sec. 14.

Frequently, it becomes necessary to calculate the derivative of the function $\varphi(\alpha)$ when the indefinite integral is complicated or even cannot be written down explicitly. Inasmuch as the parameter α is independent of x, it appears plausible that in some cases it may be permissible to perform the differentiation under the integral sign, so that one can use the formula

$$\frac{d\varphi}{d\alpha} = \int_{u_0}^{u_1} \frac{\partial f(x, \alpha)}{\partial \alpha}\, dx.$$

This formula turns out to be correct if $f(x, \alpha)$ and $\partial f(x, \alpha)/\partial \alpha$ are continuous functions in both x and α. Thus, forming the difference quotient with the aid of (51-1),

$$\text{(51-2)} \qquad \frac{\varphi(\alpha + \Delta\alpha) - \varphi(\alpha)}{\Delta\alpha} = \int_{u_0}^{u_1} \frac{f(x, \alpha + \Delta\alpha) - f(x, \alpha)}{\Delta\alpha}\, dx.$$

Now the limit, as $\Delta\alpha \to 0$, of the left-hand member of (51-2) is precisely $d\varphi/d\alpha$, whereas the limit of the expression under the integral sign is $\partial f/\partial \alpha$. Hence, if it is permissible to interchange the order of integration and calculation of the limit, one has

$$\text{(51-3)} \qquad \frac{d\varphi}{d\alpha} = \int_{u_0}^{u_1} \frac{\partial f(x, \alpha)}{\partial \alpha}\, dx.$$

The restrictions imposed on the function $f(x, \alpha)$ can be shown to be sufficient to justify the inversion of the order of these operations.

Suppose next that the limits of integration u_1 and u_0 are functions of the parameter α, so that

$$\varphi(\alpha) = \int_{u_0(\alpha)}^{u_1(\alpha)} f(x, \alpha)\, dx.$$

In this case, one can proceed as follows: Let

$$\int f(x, \alpha)\, dx = F(x, \alpha)$$

so that

$$\text{(51-4)} \qquad \frac{\partial F}{\partial x} = f(x, \alpha).$$

Then,

$$\text{(51-5)} \qquad \varphi(\alpha) = \int_{u_0(\alpha)}^{u_1(\alpha)} f(x, \alpha)\, dx = F(x, \alpha)\Big|_{x=u_0(\alpha)}^{x=u_1(\alpha)} = F(u_1, \alpha) - F(u_0, \alpha).$$

Assuming the continuity of all the derivatives involved, one can write*

$$\frac{d\varphi}{d\alpha} = \frac{\partial F(u_1, \alpha)}{\partial u_1}\frac{du_1}{d\alpha} + \frac{\partial F(u_1, \alpha)}{\partial \alpha} - \frac{\partial F(u_0, \alpha)}{\partial u_0}\frac{du_0}{d\alpha} - \frac{\partial F(u_0, \alpha)}{\partial \alpha},$$

which, upon making use of (51-4) and (51-5), becomes

$$\begin{aligned}\frac{d\varphi}{d\alpha} &= f(u_1, \alpha)\frac{du_1}{d\alpha} - f(u_0, \alpha)\frac{du_0}{d\alpha} + \frac{\partial}{\partial\alpha}[F(u_1, \alpha) - F(u_0, \alpha)] \\ &= f(u_1, \alpha)\frac{du_1}{d\alpha} - f(u_0, \alpha)\frac{du_0}{d\alpha} + \frac{\partial}{\partial\alpha}\int_{u_0(\alpha)}^{u_1(\alpha)} f(x, \alpha)\,dx.\end{aligned}$$

The partial derivative appearing in this expression means that the differentiation is to be performed with respect to α, treating u_0 and u_1 as constants. Hence, making use of (51-3),

$$\text{(51-6)}\quad \frac{d\varphi}{d\alpha} = f(u_1, \alpha)\frac{du_1}{d\alpha} - f(u_0, \alpha)\frac{du_0}{d\alpha} + \int_{u_0(\alpha)}^{u_1(\alpha)} \frac{\partial f(x, \alpha)}{\partial \alpha}\,dx.$$

This formula is known as the formula of Leibnitz, and it specializes to (51-3) when u_1 and u_0 are independent of α. The validity of this formula can be established under somewhat less restrictive hypotheses,† but the limitations imposed on the function $f(x, \alpha)$ in the foregoing discussion are usually met in problems arising in applied mathematics.

Example 1. Find $\frac{d\varphi}{dx}$, if $\varphi(\alpha) = \int_{-\alpha^2}^{2\alpha} e^{-\frac{x^2}{\alpha^2}}\,dx.$

Then

$$\begin{aligned}\frac{d\varphi}{d\alpha} &= \int_{-\alpha^2}^{2\alpha} \frac{2x^2}{\alpha^3} e^{-\frac{x^2}{\alpha^2}}\,dx - e^{-\alpha^2}(-2\alpha) + e^{-4}(2) \\ &= \int_{-\alpha^2}^{2\alpha} \frac{2x^2}{\alpha^3} e^{-\frac{x^2}{\alpha^2}}\,dx + 2\alpha e^{-\alpha^2} + 2e^{-4}.\end{aligned}$$

Example 2. Formula (51-3) is frequently used for evaluating definite integrals. Thus, if

$$\varphi(\alpha) = \int_0^\pi \log(1 + \alpha\cos x)\,dx,$$

* See Sec. 39.

† See Sokolnikoff, I. S., Advanced Calculus, Sec. 39, p. 121.

then

$$\varphi'(\alpha) = \int_0^\pi \frac{\cos x}{1 + \alpha \cos x}\, dx = \frac{1}{\alpha}\int_0^\pi \left(1 - \frac{1}{1 + \alpha \cos x}\right) dx$$
$$= \frac{1}{\alpha}\left(x + \frac{1}{\sqrt{1 - \alpha^2}} \sin^{-1} \frac{\alpha + \cos x}{1 + \alpha \cos x}\right)\Bigg|_0^\pi$$
$$= \frac{1}{\alpha}\left[\pi + \frac{1}{\sqrt{1 - \alpha^2}}\left(\sin^{-1}\frac{\alpha - 1}{1 - \alpha} - \sin^{-1}\frac{\alpha + 1}{1 + \alpha}\right)\right]$$
$$= \frac{1}{\alpha}\left(\pi + \frac{-\pi}{\sqrt{1 - \alpha^2}}\right) = \frac{\pi}{\alpha} - \frac{\pi}{\alpha\sqrt{1 - \alpha^2}}.$$

Therefore,

$$\varphi(\alpha) = \pi \int \left(\frac{1}{\alpha} - \frac{1}{\alpha\sqrt{1 - \alpha^2}}\right) d\alpha$$
$$= \pi\left(\log \alpha + \log \frac{1 + \sqrt{1 - \alpha^2}}{\alpha}\right) + c$$

or

$$\varphi(\alpha) = \pi \log (1 + \sqrt{1 - \alpha^2}) + c.$$

But, when $\alpha = 0$,

$$\varphi(0) = \int_0^\pi \log 1\, dx = 0.$$

Hence,

$$0 = \pi \log 2 + c \qquad \text{and} \qquad c = -\pi \log 2,$$

and

$$\varphi(\alpha) = \pi \log \left(\frac{1 + \sqrt{1 - \alpha^2}}{2}\right).$$

PROBLEMS

1. Find $d\varphi/d\alpha$ if $\varphi(\alpha) = \int_0^{\pi/2} \sin \alpha x\, dx$ by using the Leibnitz formula, and check your result by direct calculation.

2. Find $d\varphi/d\alpha$, if $\varphi(\alpha) = \int_0^\pi (1 - \alpha \cos x)^2\, dx$.

3. Find $d\varphi/d\alpha$, if $\varphi(\alpha) = \int_0^{\alpha^2} \tan^{-1} \frac{x}{\alpha^2}\, dx$.

4. Find $d\varphi/d\alpha$, if $\varphi(\alpha) = \int_0^\alpha \tan (x - \alpha)\, dx$.

5. Find $d\varphi/dx$, if $\varphi(x) = \int_0^{x^2} \sqrt{x}\, dx$.

6. Differentiate under the sign and thus evaluate $\int_0^\pi \frac{dx}{(\alpha - \cos x)^2}$ by using $\int_0^\pi \frac{dx}{\alpha - \cos x} = \frac{\pi}{(\alpha^2 - 1)^{1/2}}$, if $\alpha^2 > 1$.

7. Show that

$$\int_0^\pi \log (1 - 2\alpha \cos x + \alpha^2)\, dx = 0, \qquad \text{if } \alpha^2 \leq 1$$
$$= \pi \log \alpha^2, \qquad \text{if } \alpha^2 \geq 1.$$

8. Verify that

$$y = \frac{1}{k}\int_0^x f(\alpha) \sin k(x - \alpha)\, d\alpha$$

is a solution of the differential equation

$$\frac{d^2y}{dx^2} + k^2y = f(x),$$

where k is a constant.

52. Definition and Evaluation of the Double Integral. The double integral is defined and geometrically interpreted in a manner entirely analogous to that sketched above for the simple integral. Let $f(x, y)$ be a continuous and single-valued function within a region R (Fig. 38), bounded by a closed curve C, and upon the boundary C. Let the region R be subdivided in any manner into n subregions $\Delta R_1, \Delta R_2, \cdots, \Delta R_n$ of areas $\Delta A_1, \Delta A_2, \cdots, \Delta A_n$. Let (ξ_i, η_i) be any point in the subregion ΔR_i, and form the sum

$$\sum_{i=1}^{n} f(\xi_i, \eta_i)\, \Delta A_i.$$

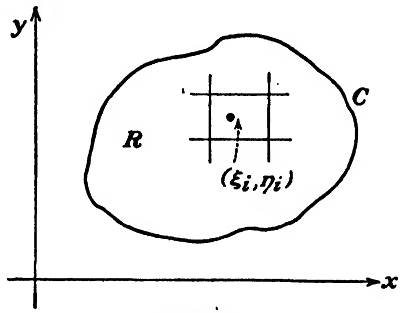

FIG. 38.

The limit of this sum, as $n \to \infty$ and all $\Delta A_i \to 0$, is defined as the double integral of $f(x, y)$ over the region R. Thus,

$$\lim_{n \to \infty} \sum_{i=1}^{n} f(\xi_i, \eta_i)\, \Delta A_i \equiv \int_R f(x, y)\, dA. \tag{52-1}$$

The region R is called the region of integration, corresponding to the interval of integration (a, b) in the case of the simple integral. The integral (52-1) is sometimes written as

$$\iint_R f(x, y)\, dx\, dy.$$

In order to evaluate the double integral, it will be simpler to consider first the case in which the region R (Fig. 39) is a rectangle bounded by the lines $x = a$, $x = b$, $y = c$, $y = d$. The extension to other types of regions will be indicated later. Subdivide R into mn rectangles by drawing the lines $x = x_1$, $x = x_2, \cdots, x = x_{n-1}$, $y = y_1$, $y = y_2, \cdots, y = y_{m-1}$ Define $\Delta x_i \equiv x_i - x_{i-1}$, where $x_0 = a$ and $x_n = b$, and define $\Delta y_j \equiv y_j - y_{j-1}$, where $y_0 = c$ and $y_m = d$. Let ΔR_{ij} be the rectangle bounded by the lines $x = x_{i-1}$, $x = x_i$, $y = y_{j-1}$, $y = y_j$. Then, if the area of ΔR_{ij} is denoted by ΔA_{ij},

$$\Delta A_{ij} = \Delta x_i\, \Delta y_j.$$

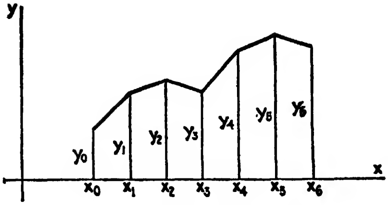

FIG. 157.

Formula (167-3) is known as the trapezoidal rule, for it gives the value of the sum of the areas of the n trapezoids whose bases are the ordinates $y_0, y_1, y_2, \cdots, y_n$. Figure 157 shows the six trapezoids in the case of $n = 6$.

If $m = 2$, (167-1) becomes

$$\begin{aligned}\int_0^2 y\,dX &= \int_0^2 \left[y_0 + X\,\Delta y_0 + \frac{(X^2 - X)}{2}\,\Delta^2 y_0\right] dX\\ &= 2y_0 + 2\,\Delta y_0 + \frac{1}{2}\left(\frac{8}{3} - 2\right)\Delta^2 y_0\\ &= 2y_0 + 2(y_1 - y_0) + \frac{1}{3}(y_2 - 2y_1 + y_0)\\ &= \frac{1}{3}y_0 + \frac{4}{3}y_1 + \frac{1}{3}y_2,\end{aligned}$$

or

$$(167\text{-}4)\qquad \int_{x_0}^{x_2} y\,dx = \frac{d}{3}(y_0 + 4y_1 + y_2).$$

Suppose that there are $n + 1$ pairs of given values, where n is even. If these $n + 1$ pairs are divided into the groups of three pairs with abscissas $x_{2i}, x_{2i+1}, x_{2i+2}$, $\left(i = 0, 1, \cdots, \frac{n-1}{2}\right)$, then (167-4) can be applied to each group. Hence,

$$\begin{aligned}(167\text{-}5)\qquad \int_{x_0}^{x_n} y\,dx &= \int_{x_0}^{x_2} y\,dx + \int_{x_2}^{x_4} y\,dx + \cdots + \int_{x_{n-2}}^{x_n} y\,dx\\ &= \frac{d}{3}(y_0 + 4y_1 + y_2) + \frac{d}{3}(y_2 + 4y_3 + y_4)\\ &\qquad + \cdots + \frac{d}{3}(y_{n-2} + 4y_{n-1} + y_n)\\ &= \frac{d}{3}[y_0 + y_n + 4(y_1 + y_3 + \cdots + y_{n-1})\\ &\qquad + 2(y_2 + y_4 + \cdots + y_{n-2})].\end{aligned}$$

Formula (167-5) is known as Simpson's rule with $m = 2$. Interpreted geometrically, it gives the value of the sum of the areas under the second-degree parabolas that have been passed through the points (x_{2i}, y_{2i}), (x_{2i+1}, y_{2i+1}), and (x_{2i+2}, y_{2i+2}), $[i = 0, 1, 2, \cdots, (n-1)/2]$.